

Article

# What Can You Do with 100 kWh? A Longitudinal Study of Using an Interactive Energy Comparison Tool to Increase Energy Awareness <sup>+</sup>

Björn Hedin <sup>1,\*</sup> <sup>D</sup> and Jorge Luis Zapico <sup>2</sup>

- Media Technology and Interaction Design, KTH Royal Institute of Technology, 10044 Stockholm, Sweden
- 2 Department of Computer Science and Media Technology, Linnaeus University, 35195 Växjö, Sweden; jorgeluis.zapico@lnu.se
- \* Correspondence: bjornh@kth.se; Tel.: +46-8-7906-990
- This is an expanded version of a conference paper. In Proceedings of the PERSUASIVE 2017: Persuasive + Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors, Amsterdam, The Netherlands, 4–6 April 2017.

Received: 31 May 2018; Accepted: 28 June 2018; Published: 2 July 2018



Abstract: Reducing the use of energy is important for several reasons, such as saving money and reducing impact on the climate. However, the awareness among non-experts of how much energy is required by different activities and appliances is generally low, which can lead to wrong prioritizations. In this study, we have developed an interactive tool to increase "energy awareness", and performed a longitudinal study to evaluate its effect. A group of 58 students first did a test to benchmark their current energy awareness, where their current knowledge of energy used for 14 different activities, such as driving vehicles and using home appliances, was measured. They then tried the interactive learning tool for 10 min. Next, they did the same test immediately after trying the tool, then again one week after trying the tool, and finally again six months after trying the tool. The results showed a significant learning effect in energy awareness with a "huge" effect size of 2.25 immediately after the intervention, a "very large" effect size of 1.70 after one week, and a "large" effect size of 0.93 after six months. The results further showed that the respondents consistently underestimated what 100 kWh could be used for, and especially so for appliances and activities requiring little energy. Before the intervention, on average they underestimated how much 100 kWh could be used for by 95.2%, and six months after the intervention the underestimation was 86.8%.

Keywords: energy awareness; technology enhanced learning; sustainable HCI; e-learning; energy conservation; energy literacy

## 1. Introduction

Energy use is an important topic for sustainability and climate change. While energy efficiencies can be created by optimization and dematerialization, promoting conservation behavior remains an important part of saving energy and avoiding rebound effects [1]. There has been much research during the last decade on using persuasive technologies and eco-visualizations to promote energy conservation. Energy conservation is a dominant topic in using persuasive technologies for sustainability and sustainable HCI (Human Computer Interaction) research [2]. Technologies such as smart meters allow users to measure and provide accurate feedback on energy consumption at both household level and appliance level. Different research activities have used this data with a persuasive intent to promote energy conservation. Many of the smart meter technologies, visualizations and applications created in this research area used quantitative data of energy using watt hour (Wh) or kilowatt hour (kWh) as the main unit [3–8].

While kWh is a familiar term for everyone paying electricity bills or selecting an appliance, it is a fairly abstract unit that may be difficult to relate to, as explored by Karjalainen [9]. Even if abstract, it is still a relevant unit. Wood and Newborough have argued that "The kWh is already familiar to most consumers and although few understand this unit, thorough comprehension is not necessary for an effective display" [10]. However, this lack of understanding of qualitative information is a part of a general lack of energy literacy [11,12]. While users know that using a lightbulb all the time, or charging their mobile phones requires energy, the lack of energy literacy or awareness needed to understand the differences in scale may lead to lack of or suboptimal conservation behavior, for example by removing a mobile phone charger during the night in order to conserve energy, but keeping underfloor heating on during the summer.

This article explores the understanding among consumers of energy, and the use of simulation as a way of increasing awareness. A prototype called "KiloWhat?" was created that uses simulation and play to practice and learn about energy consumption, exploring the following research questions:

- RQ1: What is the learning effect of using such a tool?
- RQ2: Was the energy use under- or overestimated before and after the intervention?

An experiment was performed with 58 engineering students. The experiment contained five steps (see Method section for more extensive description):

- (1) An assessment of their current energy understanding of kWh information as a baseline.
- (2) Using and experimenting with the KiloWhat prototype for 10 min.
- (3) A second test immediately after using the tool to see how the level of knowledge had changed.
- (4) A third test one week later to see how much knowledge had been lost.
- (5) A final test six months later to see how much knowledge had been lost.

This article presents the design of the KiloWhat prototype, the method and results from the test, and discusses the implications of the results for designing education services concerning energy consumption.

## 2. Design of the Prototype

The prototype is an online service available at http://kilowh.at/. The main goals of the design are to:

- (1) Make quantitative kWh information easier to relate to by providing a learning experience where the users can translate kWh into everyday activities;
- (2) Help users to learn differences in scale between the energy consumption of different activities by allowing the users to play and compare between them.

The prototype is based on the code and idea of a similar prototype named "carbon.to" available at http://carbon.kilowh.at/, which explored the same concept for increasing greenhouse gas literacy [13]. The prototype provides the possibility of exploring and comparing the energy use of 16 different activities such as driving an electric car or using a television set. The activities are divided into five different categories. The values provided were average values gathered from a search in research databases and should only be seen as rough estimates since there can be large differences within one category.

- (1) Energy generation: hours of solar panels, kg of coal, hours running in a treadmill.
- (2) Home and appliances: washing machine loads, hours with a light-emitting diode (LED) lamp on, hours with a incandescent light bulb on, hours with a fridge on, hours with Wi-Fi on, mobile phone charges, hours watching television.
- (3) Transportation: km driving a gasoline car, km driving an electric car, km driving an electric assisted bike.

- (4) Heating: hours heating a house with electric heaters, hours heating a house with geothermal pump. This category was later removed from the analysis since we found out that some participants interpreted it as averages over a year while other interpreted it as having a radiator turned on at maximum effect all the time.
- (5) Food: hamburgers (energy needed to produce).

The questions are presented in Appendix A and the reference values used are listed in Appendix B. When accessing the site, it shows by default 1 kWh compared to one of the activities selected randomly. The users can change the unit in the right side for comparing 1 kWh to other activities and increase or decrease the amounts using the plus and minus buttons on the top-right corners. The users can also change the unit on the left side to compare different activities against each other or to see what would be needed to generate that amount of energy. The interface is shown in Figure 1.



Figure 1. The interface of the kilowh.at prototype.

The main learning strategy used in this prototype is simulation, "enabling users to observe immediately the link between cause and effect" [14]. The intent is that by comparing kilowatt hours against different everyday units, and those units against each other, the users can simulate the energy use of those everyday actions and gain knowledge about them.

The software is a fork of the aforementioned carbon.to. It is developed using Ruby on Rails and JavaScript, the code is available as open source in GitHub (https://github.com/zapico/kilowhat).

As a first design iteration, the prototype was tested with 20 users. The users tested the site and provided feedback. From this feedback, several improvements were added, including a better grouping of activities using colored categories, a unification of time units to hours and the inclusion of several extra activities.

#### 3. Method

In order to evaluate the effect of the prototype, an experiment was designed. The respondents were 58 first-year engineering students who had recently started a five-year educational program within media technology. As part of an introductory course module about "Introduction to scientific research" they were required to participate in one scientific experiment. Out of the 58 students, 48 participated fully in the five parts of this experiment. In order to check which students had participated, the participation was not anonymous, but the students were informed that the email addresses collected would not be used to identify individuals. The students were also informed that the results would be used for research, and that they would be presented with the resulting paper with comments as a part of the course module. The only requirement was to participate; their actual results on the tests did not in any way affect the respondents grades or credits.

The respondents were given instructions by email, and they could do the assignment anywhere and anytime within the given time window. The email told the students about the aim of the prototype, "to increase energy awareness", an introduction to why increasing energy awareness is important, and the setup of the study.

Their first assignment consisted of three tasks that they were required to complete in a sequence within a window of two days, and the expected total time requirement was 30 min. The first task was to answer a questionnaire where they should give their best guess what 100 kWh corresponded to for 16 different activities, for example "How many kilometers riding with an electric bike", "How many hours powering a TV" or "How many charges of a mobile phone" (see Appendix A). In order to help the students understand roughly how much 100 kWh corresponds to, they were also informed that the price of 100 kWh electricity is roughly 100 Swedish crowns or about €10. The coincidence that 1 kWh roughly equals 1 Swedish crown makes it easy for users to assess if an estimate is reasonable, since money is relatively easy to relate to. They were instructed not to look up the answer anywhere while answering the questionnaire. The second task was to try out the prototype for exactly 10 min after watching a short screencast introducing the prototype. The third task was then to answer a copy of the first questionnaire. After one more week and then again after six months they were required to answer the same questionnaire in order to evaluate how much of the learning effect remained.

The first part of the analysis was to evaluate the learning effect. After receiving all responses to the questionnaire, the answers were edited [15], where the respondents whom we had good reasons to believe did not follow the instructions were removed from the analysis. Out of the 48 respondents who had answered all four questionnaires, four were removed since the log files showed that they had spent less than 10 min in trying out the prototype. These were removed from the analysis. Furthermore, we found three respondents whom for several questions had answered incorrectly by a factor of exactly 100, and their average results were also off by a factor close to 100. This led us to believe they had answered what 1 kWh would have corresponded to instead of 100 kWh. However, these were still included in the analysis, since they had used the prototype as intended but misunderstood how it worked, thereby actually decreasing their energy awareness rather than increasing it.

The answers from the remaining 44 respondents were then analyzed where both *p*-value and effect size of the learning were calculated.

The first relevant measure was the order of magnitude that the responses were distant from the correct value i.e., a response 10 times higher than the correct response is as good or bad as a response 10 times lower. Therefore, a transformation was performed on the responses, where the values of the student response was transformed by the function  $abs(log_{10}(X/C))$  where X was the student's answer and C was the "correct" reference answer. For example, if the correct answer was 10 h, then a student answer of both 5 or 20 h would have yielded a value of 0.30; if the student's answer was either 1 or 100 h both would have yielded a value of 1. Finally, the results of all 14 questions were averaged for each student. The resulting distributions were tested for normality using the Anderson–Darling normality test and were found to be roughly normally distributed and, therefore, paired *t*-tests could be used for analysis.

The second relevant measure, used to determine if the respondents underestimated or overestimated the energy use of the appliances and activities, was the same function but without using the absolute value. That means that in the examples above, an underestimation of the energy use would result in values of 0.30 and 1 respectively, but an overestimation of the energy use would have resulted in the values -0.30 and -1 respectively.

## 4. Results

The results of the two research questions are accounted for below.

## 4.1. RQ1: What Is the Learning Effect of Using Such a Tool?

The results from the first questionnaire (T0) showed that the energy awareness was low, as expected. The respondents' estimates were mostly incorrect for the activities that consumed least energy, such as Wi-Fi and charging mobile phones. The average estimation among the 616 data points generated by the 44 students who answered all four questionnaires as intended was 30.0 times higher or lower than the correct value (calculated as  $10^{\text{avg}(\text{abs}(\log 10(X/C)))})$ ). Their estimates were improved substantially in the questionnaire answered immediately after trying the prototype (T1), where the average estimate was instead only 2.62 times higher or lower than the correct value. A paired-samples *t*-test on the average score of each individual showed that scores were significantly better for test T1 immediately after trying the prototype (M = 0.419, SD = 0.416) than for the baseline test T0 before trying the prototype (M = 1.476, SD = 0.493), t(43) = 11.277, *p* < 0.001, Cohen's d = 2.248 which corresponds to a "huge" effect size [16].

The results after one more week (T2) was on average 5.66 times higher/lower than the reference value, which a significant drop compared to the results of the second questionnaire, but a paired-samples *t*-test showed that the scores still were significantly higher at test T2 (M = 0.753, SD = 0.480) than at test T0 before trying the prototype (M = 1.476, SD = 0.493), t(43) = 8.014, p < 0.001, Cohen's d = 1.697, which corresponds to a "very large" effect size [16].

Finally, after six months (T3) the average results were 11.87 times higher/lower than the reference value, and a paired-samples t-test showed that the scores were still significantly higher (M = 1.074, SD = 0.437) than at test T0 before trying the prototype (M = 1.451, SD = 0.505), t(43) = 4.590, p < 0.001, Cohen's d = 0.929 which corresponds to a large effect size [17].

These statistics are presented in Table 1.

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	AVG
T0	MEAN	0.93	1.52	1.80	1.91	1.47	1.67	2.29	1.92	1.37	0.98	0.89	0.91	2.41	0.60	1.48
	STDEV	0.62	0.77	0.96	0.83	0.90	0.83	0.89	0.74	0.79	0.61	0.51	0.51	0.97	0.51	0.49
T1	MEAN	0.33	0.38	0.46	0.49	0.44	0.56	0.46	0.38	0.34	0.30	0.30	0.53	0.57	0.32	0.42
	STDEV	0.48	0.56	0.64	0.63	0.58	0.54	0.61	0.58	0.49	0.49	0.43	0.63	0.61	0.41	0.42
T2	MEAN	0.73	0.73	0.71	0.98	0.76	0.81	0.92	0.92	0.58	0.69	0.58	0.65	1.01	0.46	0.75
	STDEV	0.66	0.64	0.89	0.87	0.67	0.71	0.82	0.75	0.58	0.57	0.49	0.67	0.90	0.69	0.48
T3	MEAN	0.71	1.30	1.14	1.47	1.02	1.07	1.50	1.47	0.91	0.80	0.71	0.70	1.76	0.48	1.07
	STDEV	0.50	0.83	0.86	0.76	0.91	0.71	0.87	0.80	0.57	0.60	0.43	0.45	0.97	0.55	0.44
T1-T0	M0-M1	0.60	1.14	1.34	1.42	1.03	1.11	1.84	1.54	1.03	0.68	0.59	0.38	1.83	0.29	1.06
	t	5.38	8.30	7.20	9.68	6.79	8.50	11.62	11.89	7.51	5.24	5.50	2.98	11.14	2.56	11.28
	р	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	Cohen's d	1.03	1.56	1.30	1.86	1.23	1.56	2.15	2.31	1.35	1.02	1.08	0.71	2.02	0.50	2.25
T2-T0	M0-M2	0.20	0.79	1.09	0.93	0.70	0.86	1.37	0.99	0.79	0.29	0.31	0.26	1.40	0.15	0.72
	t	1.59	5.26	6.01	6.00	5.00	6.57	9.03	6.37	5.57	2.14	3.19	2.17	7.28	1.06	8.01
	р	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.04	0.00	0.30	0.00
	Cohen's d	0.35	1.03	1.24	1.32	0.95	1.31	1.87	1.37	1.04	0.44	0.67	0.54	1.51	0.27	1.70
T3-T0	M0-M3	0.22	0.23	0.66	0.44	0.44	0.61	0.79	0.44	0.46	0.18	0.18	0.21	0.64	0.13	0.40
	t	2.17	1.46	4.25	2.86	2.67	4.11	4.32	3.08	3.73	1.47	1.88	1.97	3.62	1.20	4.59
	р	0.04	0.15	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.15	0.07	0.05	0.00	0.24	0.00
	Cohen's d	0.42	0.32	0.86	0.59	0.57	0.82	0.91	0.68	0.70	0.31	0.37	0.40	0.77	0.27	0.93

Table 1. Statistical analysis of the four tests T0–T4 for each of the 14 questions.

#### 4.2. RQ2: Was the Energy Use Under- or Overestimated before and after the Intervention?

The results above used the absolute value of the deviations from the correct reference value, which is suitable for discussing the learning effect since an overestimation of a value should be considered equally incorrect as an underestimation. From a sustainability point of view, however, it is also important to know if respondents underestimate or overestimate how much energy appliances and activities use, both as a baseline (i.e., the energy awareness without an intervention) and to see if this possible bias can partially or fully be removed by using a tool such as described in this study.

When taking into account the direction (i.e., underestimation or overestimation) of the answers provided by the respondents, the results from the first questionnaire (T0) showed that what the 100 kWh could be used for was severely underestimated. The average value of the estimations (calculated as  $avg(log_{10}(X/C)))$  was -1.319 as can be seen in the row MEAN(T0) in Table 2. This corresponds to the fact that they underestimated how much the 100 kWh could be used for by, on average, 95.2%.

The results from the questionnaire answered immediately after using the prototype (T1) showed a significant improvement in energy awareness (t(43) = 9.664, p < 0.001, Cohen's d = -1.865), which corresponds to a very large effect size [16]. The average value of the estimations (calculated as avg(log<sub>10</sub>(X/C))) had increased from -1.302 to -0.191. This corresponds to that they now underestimated how much the 100 kWh could be used for on an average by 35.6% as compared to the previous 95.2%.

The results from the questionnaire answered one week after using the prototype (T2) also showed a significant improvement in energy awareness compared to T0 (t(43) = 7.434, p < 0.001, Cohen's d = -1.615), which corresponds to a very large effect size [16]. The average value of the estimations (calculated as avg(log<sub>10</sub>(X/C))) was now -0.373 which was an increase from T0 (-1.319) but a decrease from T1 (-0.191) as could be expected. This corresponds to them now underestimating how much the 100 kWh could be used for on an average by 57.6%.

Finally, the results from the questionnaire answered six months after using the prototype (T3) also showed a significant improvement in energy awareness compared to T0 (t(43) = 3.624, p < 0.001, Cohen's d = -0.753), which a medium effect size [17]. The average value of the estimations (calculated as avg(log<sub>10</sub>(X/C))) was now -0.878 which was an increase from T0 (-1.319) but a decrease from both T1 (-0.191) and T2 (-0.373) as could be expected. This corresponds to them now underestimating how much the 100 kWh could be used for on average by 86.8%. The results are shown in Table 2.

Finally, looking into more detail into the 14 different use cases, the use cases where the 100 kWh would provide "a lot" of the unit in question (hours, kilometer or pieces) were consistently underestimated more than where it would provide "a little". For example, Q7–Q9 asked about how many kilometers an electric bike (Q7), an electric car (Q8) or a petrol-driven car (Q9) could travel using 100 kWh, and in all four questionnaires the distance that could be traveled by an electric bike was underestimated the most, followed by the electric car, which in turn was followed by the petrol-driven car.

Two different types of correlations were calculated. The first was the Pearson correlation of the averages of  $\log_{10}(X/C)$  for the students' answers to each question, with  $\log_{10}(C)$  for the corresponding question for each of the four questionnaires. The second was the Pearson correlation for Cohen's d (the effect size) for each of the questions with  $\log_{10}(C)$  for the corresponding question for each of the four questions with  $\log_{10}(C)$  for the corresponding question for each of the four questions with  $\log_{10}(C)$  for the corresponding question for each of the four questions with  $\log_{10}(C)$  for the corresponding question for each of the four questions with  $\log_{10}(C)$  for the corresponding question for each of the four question for each of

The results in Table 3 shows a very strong negative correlation for all calculations. This means that the respondents not only underestimated most results even after the intervention, but also that the level of underestimation is strongly inversely correlated with the energy use required, or to put it in other words that the energy use of appliances and activities that require little energy is overestimated. However, the learning effect after trying the prototype was also higher for these appliances or activities as can be seen in the strong inverse correlation with the effect sizes.

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	AVG
Т0	MEAN	-0.50	-1.41	-1.78	-1.89	-1.44	-1.63	-2.29	-1.88	-1.35	-0.87	-0.50	-0.83	-2.41	0.30	-1.32
	STDEV	1.00	0.96	1.01	0.86	0.94	0.92	0.89	0.83	0.83	0.76	0.90	0.64	0.97	0.73	0.62
T1	MEAN	-0.04	-0.23	-0.37	-0.39	-0.27	-0.28	-0.39	-0.22	-0.10	-0.08	0.00	0.14	-0.41	0.00	-0.19
	STDEV	0.58	0.64	0.70	0.70	0.68	0.73	0.65	0.66	0.59	0.57	0.52	0.81	0.73	0.52	0.49
T2	MEAN	-0.08	-0.42	-0.62	-0.88	-0.36	-0.43	-0.85	-0.78	-0.17	-0.05	-0.03	0.16	-0.88	0.17	-0.37
	STDEV	0.99	0.88	0.96	0.97	0.96	0.99	0.90	0.91	0.81	0.90	0.77	0.92	1.03	0.81	0.66
T3	MEAN	-0.34	-1.15	-1.04	-1.46	-0.88	-0.99	-1.50	-1.43	-0.78	-0.51	-0.30	-0.41	-1.68	0.18	-0.88
	STDEV	0.80	1.03	0.99	0.78	1.05	0.81	0.88	0.87	0.74	0.87	0.78	0.73	1.10	0.71	0.59
T1-T0	M0-M1	-0.46	-1.18	-1.41	-1.50	-1.17	-1.35	-1.90	-1.66	-1.24	-0.79	-0.50	-0.96	-1.99	0.30	-1.13
	t	2.61	7.09	7.45	9.84	7.03	8.85	11.96	10.98	8.07	4.98	3.20	6.27	11.64	2.38	9.66
	р	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
	Cohen's d	-0.46	-1.30	-1.37	-1.93	-1.32	-1.72	-2.27	-2.14	-1.51	-0.94	-0.56	-1.53	-2.23	0.44	-1.87
T2-T0	M0-M2	-0.42	-1.00	-1.15	-1.01	-1.08	-1.20	-1.44	-1.11	-1.18	-0.82	-0.46	-0.99	-1.52	0.13	-0.95
	t	2.17	5.02	5.94	5.90	6.62	6.95	9.03	6.28	6.55	4.69	2.80	5.51	7.47	0.81	7.34
	p	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.42	0.00
	Cohen's d	-0.46	-1.03	-1.24	-1.34	-1.43	-1.55	-1.95	-1.41	-1.39	-1.10	-0.55	-1.45	-1.65	0.18	-1.62
T3-T0	M0-M3	-0.16	-0.27	-0.74	-0.43	-0.56	-0.63	-0.80	-0.45	-0.57	-0.36	-0.20	-0.41	-0.72	0.13	-0.44
	t	0.88	1.36	4.24	2.68	3.13	3.66	4.29	2.84	3.92	2.15	1.31	2.79	4.17	0.89	3.62
	p	0.38	0.18	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.04	0.20	0.01	0.00	0.38	0.00
	Cohen's d	-0.17	-0.30	-0.90	-0.55	-0.71	-0.74	-0.91	-0.62	-0.80	-0.49	-0.26	-0.64	-0.96	0.19	-0.75

**Table 2.** Results of paired *t*-tests for the three different questionnaires.

**Table 3.** Pearson correlations for all 14 questions between  $log_{10}(C)$  and with (1) the average replies for  $log_{10}(X/C)$  for all questions in questionnaires T0–T4, and (2) between  $log_{10}(C)$  and the learning effect expressed as Cohen's d.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	CORREL
С	67	1000	16,667	4488	2500	1493	17,544	5000	472	125	47	104	18,349	26	
Log10(C)	1.826	3.000	4.222	3.652	3.398	3.174	4.244	3.699	2.674	2.097	1.672	2.017	4.264	1.415	
T0 Average	-0.503	-1.413	-1.779	-1.890	-1.440	-1.627	-2.292	-1.881	-1.349	-0.871	-0.497	-0.826	-2.406	0.304	-0.953
T1 Average	-0.043	-0.235	-0.370	-0.394	-0.269	-0.282	-0.391	-0.219	-0.104	-0.083	-0.001	0.138	-0.414	0.000	-0.921
T2 Average	-0.083	-0.416	-0.625	-0.880	-0.356	-0.429	-0.849	-0.775	-0.168	-0.054	-0.033	0.160	-0.884	0.170	-0.936
T3 Average	-0.342	-1.146	-1.038	-1.460	-0.884	-0.992	-1.496	-1.433	-0.780	-0.509	-0.295	-0.415	-1.681	0.176	-0.929
T0–T1 Cohen's d	-0.455	-1.296	-1.369	-1.929	-1.320	-1.723	-2.270	-2.140	-1.505	-0.942	-0.558	-1.529	-2.229	0.443	-0.832
T0–T2 Cohen's d	-0.460	-1.032	-1.240	-1.345	-1.430	-1.551	-1.951	-1.409	-1.386	-1.098	-0.554	-1.452	-1.648	0.182	-0.772
T0–T3 Cohen's d	-0.170	-0.302	-0.895	-0.545	-0.708	-0.737	-0.914	-0.623	-0.796	-0.492	-0.262	-0.637	-0.956	0.188	-0.799

#### 5. Conclusions and Discussion

In this paper we have implemented and tested a tool for increasing energy awareness. We set up two research questions:

#### 5.1. RQ1: What Is the Learning Effect of Using Such a Tool?

The results first confirmed our view that energy awareness in general was very low. However, spending 10 min with the tool increased energy awareness significantly, with an effect size after one week of 1.697 and, more importantly, a learning effect size of 0.929 after six months. Effect sizes for learning interventions were studied by Hattie [18], who in a study of more than 800 meta analyses about learning came to the conclusion that the average learning intervention had an effect size of 0.4. Cohen classifies an effect size of at least 0.8 as large, and a learning effect size of 0.929 after six months is therefore quite substantial.

#### 5.2. RQ2: Was the Energy Use Under- or Overestimated before and after the Intervention?

We found that the respondents generally underestimated how much 100 kWh could be used for, and that this was especially the case for equipment requiring small amounts of energy like a Wi-Fi router or to charge a mobile phone. The underestimation was highly inversely correlated with the energy required by the appliance or activity, but the learning effect was also greater for these low-energy appliances/activities.

#### 5.3. General Discussion

While the focus on numbers and rational approaches to sustainable behavior has been criticized [19], increasing the capacity of understanding electricity information is an important precondition in allowing better energy decisions, and it will become more important with current trends of electric mobility and small-scale electricity production. This study confirmed that the understanding of quantitative energy information was low, even with a group of engineering students. This low energy understanding can affect energy decisions and create misguided efforts with low energy saving potential. The tool used in this study shows how providing anchor points that the users could refer to, increased their energy awareness, as having reference points to concrete everyday actions can help in having a more intuitive understanding of quantitative energy information.

The learning effect was, as shown, quite substantial even after six months, but as expected it was much higher shortly after the intervention. An increased long-term learning effect could likely be achieved if a spaced repetition [20] learning strategy was used, where the prototype was used again after, for example, one month and six months. The spacing effect, where studying efforts spread out over a longer period of time leads to better learning, has been shown to be "one of the most dependable and replicable phenomena in experimental psychology" [21]. If the society would be interested in increasing energy awareness among the population, the allocation of 3 times 10 min of high school students' time for exercises such as this would, therefore, likely have a strong long-term effect.

The tool discussed in this article does not focus on changing behavior but on the energy awareness or ability of understanding energy information. This could be seen to be more relevant for people already interested in energy issues and their behavior, while it may not be of interest for other users, as motivation varies [22]. The group used in this study could not be said to be either especially environmentally friendly, nor the opposite, and it could be expected that people with a genuine interest in environmental issues would put more effort into using the tool and thereby achieve better learning.

However, a potential risk with using a tool like this is that it could have adverse effects. Since the respondents in all four tests overestimated the energy use in most use cases, increased energy awareness could make the users realize how little energy many appliances and activities actually use, and thereby it could result in a rebound effect where the users of the system instead start using more energy than before.

## 5.4. Limitations and Future Research

One limitation of this study is that the respondents were not representative of society in general. They were young and many of them lived with parents and/or did not pay for energy use. They were also relatively well educated and as engineering students could be expected to have a better knowledge about energy and kWh than most other people in their age group.

The comparisons used in the presented tool are mostly short-term actions focused on individuals, such as using an appliance or driving a car. It could be interesting to explore the use of interactive tools for supporting the understanding of energy information in a bigger scale. This could include longer time frames (such as energy used in a year), and bigger space frames (such as energy used in a city or a country).

Future research will also include comparing this kind of learning with more conventional learning methods such as studying from a list. As we saw in our original study [23] the students in general preferred this interactive learning from a more conventional learning method, which in itself can motivate the development of tools like this, but the possible differences in learning remains to be evaluated.

Author Contributions: The kilowh.at prototype was designed and implemented by J.L.Z. The study was designed, and data was collected and analyzed by B.H. The paper was written jointly.

Funding: This research was funded by the Swedish Energy Agency.

Conflicts of Interest: The authors declare no conflict of interest.

## Appendix A. Questions Asked in T0 and "Correct" Reference Values (Translated from Swedish). Corresponding Questions Were Asked for T1–T3

"You now should estimate what 'X" is for the questions below in answer to the statement "100 kWh corresponds to ...". Roughly calculated 1 kWh corresponds to 1 Swedish krona as a comparison. Do not attempt to find the answers anywhere, it is your spontaneous answer that we are interested in.

- Q1: X hours average electricity production from 10-square-meter solar cells.
- Q2: X hours running on a treadmill.
- Q3: X hours with a LED light on.
- Q4: X hours of refrigerator use.
- Q5: X hours with a old lightbulb on.
- Q6: X hours with a normal TV on.
- Q7: X hours with a Wi-Fi router on.
- Q8: X km riding on an electric bike.
- Q9: X km riding with an electric car.
- Q10: X km riding with a petrol-driven car.
- Q11: X one kilogram bags of coal (energy generated by).
- Q12: X runs with a wash machine.
- Q13: X charges of a mobile phone.
- Q14: X hamburgers (energy required to produce).

## Appendix B

Table A1. Refe	rence values	for the 14 d	questions.
----------------	--------------	--------------	------------

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
67	1000	16,667	4488	2500	1493	17,544	5000	472	125	47	104	18,349	26

## References

- 1. Berkhout, P.H.; Muskens, J.C.; Velthuijsen, J.W. Defining the rebound effect. *Energy Policy* **2000**, *28*, 425–432. [CrossRef]
- DiSalvo, C.; Sengers, P.; Brynjarsdóttir, H. Mapping the landscape of sustainable HCI. In Proceedings of the SIGCHI Conference on ACM Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 1975–1984.
- 3. Hargreavesn, T.; Nye, M.; Burgess, J. Making energy visible: A qualitative field study of how householders interact with feedback from smart energy monitors. *Energy Policy* **2010**, *38*, 6111–6119. [CrossRef]
- 4. Hargreavesn, T.; Nye, M.; Burgess, J. Keeping energy visible? Exploring how householders interact with feedback from smart energy monitors in the longer term. *Energy Policy* **2013**, *52*, 126–134. [CrossRef]
- 5. Murtagh, N.; Nati, M.; Headley, W.R.; Gatersleben, B.; Gluhak, A.; Imran, M.I.; Uzzell, D. Individual energy use and feedback in an office setting: A field trial. *Energy Policy* **2013**, *62*, 717–728. [CrossRef]
- McCalley, T.; Kaiser, F.; Midden, C.; Keser, M.; Teunissen, M. Persuasive Appliances: Goal Priming and Behavioral Response to Product-Integrated Energy Feed-Back; IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E., Eds.; PERSUASIVE 2006, LNCS 3962; Springer: Berlin/Heidelberg, German, 2006; pp. 45–49.
- Gamberini, L.; Spagnolli, A.; Corradi, N.; Jacucci, G.; Tusa, G.; Mikkola, T.; Zamboni, L.; Hoggan, E. *Tailoring Feedback to Users' Actions in a Persuasive Game for Household Electricity Conservation*; Bang, M., Ragnemalm, E.L., Eds.; PERSUASIVE 2012, LNCS 7284; Springer: Berlin/Heidelberg, German, 2012; pp. 100–111.
- Yun, R.; Lasternas, B.; Aziz, A.; Loftness, V.; Scupelli, P.; Rowe, A.; Kothari, R.; Marion, F.; Zhao, J. *Toward the* Design of a Dashboard to Promote Environmentally Sustainable Behavior among Office Workers; Berkovsky, S., Freyne, J., Eds.; PERSUASIVE 2013, LNCS 7822; Springer: Berlin/Heidelberg, German, 2013; pp. 246–252.
- 9. Karjalainen, S. Consumer preferences for feedback on household electricity consumption. *Energy Build.* 2011, 43, 458–467. [CrossRef]
- 10. Wood, G.; Newborough, M. Energy-use information transfer for intelligent homes: Enabling energy conservation with central and local displays. *Energy Build.* **2007**, *39*, 495–503. [CrossRef]
- 11. DeWaters, J.E.; Powers, S.E. Energy literacy of secondary students in New York State (USA): A measure of knowledge, affect, and behavior. *Energy Policy* **2011**, *39*, 1699–1710. [CrossRef]
- 12. Brounen, D.; Kok, N.; Quigley, J.M. Energy literacy, awareness, and conservation behavior of residential households. *Energy Econ.* **2013**, *38*, 42–50. [CrossRef]
- 13. Zapico Lamela, J.L.; Turpeinen, M.; Guath, M. Kilograms or cups of tea: Comparing footprints for better CO<sub>2</sub> understanding. *Psychnol. J.* **2011**, *9*, 43–54.
- 14. Oinas-Kukkonen, H.; Harjumaa, M. Persuasive systems design: Key issues, process model, and system features. *Commun. Assoc. Inf. Syst.* **2009**, *24*, 28.
- 15. Cohen, L.; Manion, L.; Morrison, K. Research Methods in Education; Routledge: London, UK, 2000.
- 16. Sawilowsky, S.S. New Effect Size Rules of Thumb. J. Mod. Appl. Stat. Methods 2009, 8, 597–599. [CrossRef]
- 17. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Erlbaum Associates: Hillsdale, NJ, USA, 1988.
- 18. Hattie, J. Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement; Routledge: London, UK, 2009.
- 19. Strengers, Y. Smart energy in everyday life: Are you designing for resource man? *Interactions* **2014**, *21*, 24–31. [CrossRef]
- Hedin, B. Exploring Opportunistic Use of Mobile Devices for Studying in Higher Education. (TRITA-CSC-A). Ph.D. Thesis, KTH Royal Institute of Technology, KTH, Media Technology and Interaction Design, MID, Stockholm, Sweden, 2014.
- 21. Dempster, F.N. The spacing effect: A case study in the failure to apply the results of psychological research. *Am. Psychol.* **1988**, 43, 627–634. [CrossRef]

- 22. He, H.A.; Greenberg, S.; Huang, E.M. One size does not fit all: Applying the transtheoretical model to energy feedback technology design. In Proceedings of the SIGCHI Conference on ACM Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 927–936.
- Hedin, B.; Zapico, J. Kilowh.at—Increasing Energy Awareness Using an Interactive Energy Comparison Tool. In Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors, Proceedings of the 12th International Conference, PERSUASIVE 2017, Amsterdam, The Netherlands, 4–6 April 2017; de Vries, P.W., Oinas-Kukkonen, H., Siemons, L., Beerlage-de Jong, N., van Gemert-Pijnen, L., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 175–185.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).