Computing Educational Activities Involving People Rather Than Things Appeal More to Women (CS1 Appeal Perspective)

Melissa Høegh Marcher*
CCER: Center for Computing Education Research
IT University of Copenhagen
Copenhagen, Denmark

Ingrid Maria Christensen*
CCER: Center for Computing Education Research
IT University of Copenhagen
Copenhagen, Denmark

Paweł Grabarczyk
CCER: Center for Computing Education Research
IT University of Copenhagen
Copenhagen, Denmark

Therese Graversen
CCER: Center for Computing Education Research
IT University of Copenhagen
Copenhagen, Denmark

Claus Brabrand
CCER: Center for Computing Education Research
IT University of Copenhagen
Copenhagen, Denmark

ABSTRACT

Prior research on recruitment of women to computing has established that computing tasks involving People rather than Things have been perceived as much more appealing by female high-school students (potentially recruitable as university computing students). This paper changes the focus from prospective to current university students and presents the results of a new experiment that advances and moves beyond earlier research in two crucial respects. First of all, the participants of the experiment are N=152 university students, who already study computing, rather than general high-school students. Second of all, the choice between a People-themed versus an *isomorphic* Things-themed version of an educational task now pertains to *real* (in fact, mandatory) assignments that the students had to perform, rather than hypothetical tasks. The change of experimental context, design, and methodology allows us to complement previous findings related to *recruitment* with suggestions significant for computing educational activities. The overall findings of the new experiment are consistent with that of the previous one. We find that, also at university, there is a visible preference for choosing People themed over Things themed computing tasks amongst women. The results also expose considerable variation between tasks in the effect of gender observed. At the same time, male students, in general, seem to be either indifferent to the themes or to slightly prefer People versions. This suggests that educators should consider favoring People themed assignments over ones involving Things.

*Both authors contributed equally to this research.

ICER 2021, August 16–19, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8326-4/21/08…$15.00
https://doi.org/10.1145/3446871.3469761

CCS CONCEPTS

• **Social and professional topics → Computing education programs**; *Computer science education*; *Information technology education*; *Software engineering education*.

KEYWORDS

computing, computer science, software engineering, STEM, education, student experiments, diversity, gender

ACM Reference Format:
Melissa Høegh Marcher, Ingrid Maria Christensen, Paweł Grabarczyk, Therese Graversen, and Claus Brabrand. 2021. Computing Educational Activities Involving People Rather Than Things Appeal More to Women (CS1 Appeal Perspective). In *Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021), August 16–19, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3446871.3469761

1 INTRODUCTION

There is a high demand for information and communication technology (ICT) specialists. According to Eurostat,[1] 53% of companies in the EU reported, in 2017, that they had had difficulties filling ICT positions [18]. In 2019, the percentage of female ICT specialists was just 17.9% within the EU [19]. There is no doubt that this is a social group with the biggest potential for growth. At university, the trend is similar. According to Informatics Europe, the average percentage of bachelor degrees in informatics (computing, computer science, IT, and ICT) awarded to female students in 2019 was only 20.3% across 19 European countries [17].

The gender imbalance is, in part, due to the sparse number of women entering the field. Recruitment of more female students is an important and active research topic specifically in the context of computing [4, 37] and, more generally, in the broader context of STEM [3, 11, 12, 16, 21]. Recruitment is vital for catching up with the high demand in the industry, for diversifying the field, and for contributing to a less biased and stereotyped computing environment.

In a paper from 1998, Lippa found evidence that gender strongly relates to a preference for working with People vs Things [27]. In

[1]Eurostat: The statistical office of the European Union: https://ec.europa.eu/eurostat

three studies, a total of 2,361 people were assessed on the People–Things dimension. The studies consistently showed that women, on average, were more interested in working with People, while men were more interested in working with Things. A later paper from Lippa shares the same overall conclusion [28].

In our previous study [13] pertaining to recruitment, we applied Lippa's hypothesis in the context of computing education. We presented high-school students with hypothetical computing educational tasks in two versions—a People-themed version and a Things-themed version—and asked them to rate their preference for one over the other.

We studied the impact of the themes of educational tasks (scenarios) on how appealing they seem to be for male and female students (as proxies for potential future university students) [13]. Our research uncovered that female high-school students, in general, perceive themes involving People rather than Things as more appealing, whereas the men are more indifferent as to the particular theme of the educational tasks (although the men also, as a whole, exhibited a *slight* preference for tasks involving People over Things). In fact, women have 2.7 times higher odds, than the men, to prefer tasks involving People (p-value = 3.4×10^{-13}, based on an experiment with around 500 high-school students). Our paper recommended that recruitment activities feature educational tasks involving People themes because they appeal to women to a higher extent, without discouraging the men.

In this paper, we still consider the appeal of educational themes on gender, but the focus is shifted from *prospective* to *current* university students. Instead of looking at how to admit more women to computing, this paper will instead look at how to appeal to the ones already admitted.

There is an important ethical consideration: If the techniques proposed in our earlier work are applied to *recruitment* (emphasizing People tasks), then they should also be applied towards *retention*. Naturally, changing the way we advertise and present a given field to the public (something we suggested in our previous paper) obliges us to deliver courses that are consistent with the advertising strategy. If we fail to do that, female students choosing studies within computing as a result of attending such recruitment activities may not have their expectations met. Aside from general discontent, we speculate that this may ultimately cause higher dropout rates among women. After all, our prior research documents that high-school women, in general, have a significantly lower preference for tasks involving Things than for the ones involving People. We wanted to make the CS education more engaging, especially to women hoping that this may have a visible effect on retention.

1.1 Contributions

This paper will investigate the impact of the themes of educational tasks on gender. The paper will significantly extend upon previous research [13] and present the results of an observational study with N=152 students that:

- investigates *university* (rather than *high-school*) students;
- considers *real* (rather than *hypothetical*) assignments; *and*
- focuses on *current* (rather than *prospective*) university students.

Aside from gender, this paper will also investigate *variation* according to tasks and individuals.

2 EXPERIMENT

We outline our observational study in terms of *objectives*, *context*, *subjects*, *assignments*, *treatment*, *design*, *execution*, and *analysis*.

2.1 Objectives

This paper addresses the following three research questions:

> **RQ1:** To what extent does the gender of software development students impact the preferences for People vs Things in computing assignments?
>
> **RQ2:** To what extent do preferences vary between assignments?
>
> **RQ3:** To what extent do preferences vary between individual students?

The first research question (RQ1) addresses female and male Software Development students' *preferences* on the People–Things spectrum. The second and third research questions (RQ2 & RQ3) addresses *variation* according to *tasks* and *individuals*.

2.2 Context

For this experiment, we settled on the *Introductory Programming* (CS1) course at the IT University of Copenhagen. This is a first-semester 15 ECTS[2] mandatory course on the three-year Bachelor of Software Development. The course introduces basic programming and object-orientation using Java, BlueJ [7], along with an Object-Oriented Conceptual Framework [26] and emphasizes a strong correspondence between reality and model via object-oriented modeling of reality involving object-oriented analysis, design, and implementation. The course has been managed and co-run for the last seven years by the last author whereby we obtained permission to run this experiment.

2.3 Participants

Since the focus of this study is on *retention* (rather than *recruitment*), we settled on the first-semester university course with a recently admitted student cohort of approximately two hundred students, highly representative of the type of students educational programmes would want to retain. 37 female students and 115 male students chose to participate in the study. Please note that unlike our previous study [13], these were students who have thus already (in fact, not long before the experiment) made the "vocational choice" to study Software Development.

2.4 Assignments

During the course *Introductory Programming*, students are given four mandatory individual hand-in assignments. For this experiment, we obtained permission to re-design the last three assignments, provided that they respected the pre-existing course syllabus from the curriculum. Tables 1 & 2 provide characterizations of the

[2] One academic year is 60 ECTS (European Credit Transfer and Accumulation System).

four assignments along with the People- and Things-instantiated versions of the given scenario. All assignments were programming assignments, starting with the basics in assignment A and, for B–D, gradually progressing from modifying existing code to creating code from scratch. Excerpts from the assignments can be found in appendices.

Assignment A was not part of the experiment and thus exists in only one version. Assignment B involved *modifying* existing code, so that it adheres to object-oriented principles, in particular, inheritance. Assignment C was about *modifying* an existing program and *creating* exceptions and exception handling for it. Assignment D involved *creating* a project from scratch with a graphical user interface using the model-view-controller design pattern, optimizing cohesion, and adhering to the principles of separation of concerns.

The pre-existing course specification stipulated that students are required to get three out of four assignments formally approved to be eligible for the exam. Hence, we are not guaranteed to get an equal number of data points for each assignment.

2.5 Treatment

Each of the assignments was designed in two versions: a People-themed version and an *isomorphic* Things-themed version.

The two versions differed only with respect to their thematic instantiation but were isomorphic in all other aspects. This includes new vs old, useful vs useless, and familiar vs unfamiliar. The two versions were *visually* and *syntactically isomorphic*; i.e., the assignment descriptions were of equal length and composition and the sentences followed the same syntactic structure;[3] for example:

"A	***hotel employee***	***wants***	*to* ***have***	*a visual overview."*	[People]
"A	***screen***	***has***	*to* ***give***	*a visual overview."*	[Things]
|	|	|	|	|	
article (indefinite)	noun (singular)	verb (modal)	verb (infinitive)	noun phrase (direct object)	

```
[X] [X] [ ]
[ ] [ ] [X]
[ ] [ ] [X]
[ ] [ ] [ ]
```

Figure 1: Sample illustration from assignment C.

Any figures included in the assignment were identical in the two versions, except for any text pertaining to a specific People- or Things-instantiation. In some cases, it was possible to use the exact same figure in both versions. Figure 1 shows a figure from assignment C, which represented *booked rooms in a hotel* in the People version and *booked containers on a cargo-ship* in the Things version.

2.6 Design

Since this observational study taps into an existing course and infrastructure, the experiment design is mostly dictated by the course context (see Sections 2.2 & 2.4). For each of the three assignments (B, C, and D), the students simply had to freely choose which of the two versions of the assignment (People- or Things-version) they wanted to complete and hand in.

2.7 Execution

Pilot. Before distributing the three assignments to the students, each of the assignments were checked for technical correctness and conformity to People or Things. 1-2 teaching assistants from the course *Introductory Programming* read through the assignments for technical correctness. The course manager, who is also a co-author of this paper, checked the assignments for accordance with the course curriculum, technical correctness, and conformity to People or Things.

Experiment. The assignments were distributed via an internal learning platform in a folder containing both the People and the Things version. We made sure to deliberately alternate the alphabetical ordering of the names of the People vs the Things version so that none of the versions were consistently presented first to the students, to not enforce a bias from the order of presentation (see Section 4).

Each assignment contained an introduction explaining the guidelines for working on and handing in the assignment. The students were told that they could choose one of two versions of the assignment. The assignments had to be completed by the student alone, but they were allowed to talk to each other. Further, we told the students that the assignments were similar enough for them to be able to discuss the assignment with a peer, even if they had chosen a different version. We, importantly, stressed that they should choose based on *their own* interest. Upon completion of assignments, the students handed in using the internal learning platform from which we are able to extract data of the preferences of versions.

The assignments were introduced to the students as mandatory activities of the course and once they were all concluded, students were sent a consent form for inclusion in our study. In this connection the students also reported their gender. It was repeatedly emphasised to students that their decision on whether to participate in the study would have no bearing on the exam assessment.

2.8 Analysis Method

We analyse the binary choice made by a student for a particular assignment using a logistic regression model with mixed effects [1, 31]. An important feature of this model is that it appropriately makes use of all available data, even though some students may not have submitted all three assignments.

The model describes the odds of a student choosing the People version of the assignment over the Things version. The odds are defined as the ratio between the probability of choosing the People version and the probability of choosing the Things version.

We allow the odds of choosing the People version to be different for each combination of gender (female/male) and assignment (B–D); in particular, this means that the preferences may vary according to the specific formulation and topic of the assignment and that they may vary differently for the two genders considered. We further include in the model a Gaussian random effect (i.e., a

[3] Please note that, in Danish, "hotel employee" (People version) is *one* word, therefore not breaking syntactic isomorphism with the (one) word "screen" (Things version).

Table 1: Characterization of the mandatory hand-in assignments in the course.

ID	Week	Topic	Mode	Scenario	People-Version	Things-Version
A	4	Basic Programming	Create	Forest Growth	Forest Growth	
B	7	Inheritance	Modify	Entity Management	**Employee Management**	**Storage Management**
C	10	Exception Handling	Modify & Create	Booking System	**Hotel Booking**	**CargoShip Booking**
D	13	GUI & MVC	Create	Value Tracker	**Customer Tracker**	**Electricity Tracker**

Table 2: Quantification of the size of the assignments.

ID	#Words	#Tasks	#Subtasks	#Figures
A	902	3	18	0
B	1,005	4	13	3
C	861	4	15	1
D	1,018	3-4	20-21	2

Table 3: Number of assignment versions (People vs Things) chosen by male and female students (in absolute numbers).

	Men		Women		
ID	People	Things	People	Things	Total
A	115		37		152
B	62	53	23	13	151
C	57	55	29	8	149
D	62	46	34	3	145

random intercept) for students, which may be thought of as an inherent student-specific level for the preferences regarding People and Things. To compute the odds for a specific student choosing the People version for an assignment, this individual level is added to the overall odds (at log scale) for choosing the People version among students in the relevant gender group.

For each of the three assignments, a student has some probability of choosing the People version. According to the model this probability p, or equivalently, the log-odds ($\log\{p/(1-p)\}$) are random. This is because the log-odds are modelled as the sum of two components: 1) the overall log-odds for the relevant combination of gender and assignment and 2) the inherent student-specific addition, which is drawn from a Gaussian distribution (mean 0 and variance estimated from data) and is the same for all assignments. We may therefore consider the correlation between the students' three probabilities of choosing the People version; the model implies that this correlation is the same between any two of the three probabilities (because the random component is the same across assignments). It may be argued that data has a temporal aspect as, for instance, assignments were given in particular weeks of the course. However, we see no reason to suggest that the correlation between the student-specific probabilities should depend on the order and time between assignments.

The model is fitted using standard libraries, `lme4` and `ordinal`, in R. See Agresti and Alan [1] (Chapter 13) and Larsen et al. [23] for a detailed discussion of the interpretation of mixed models for binary data.

3 RESULTS

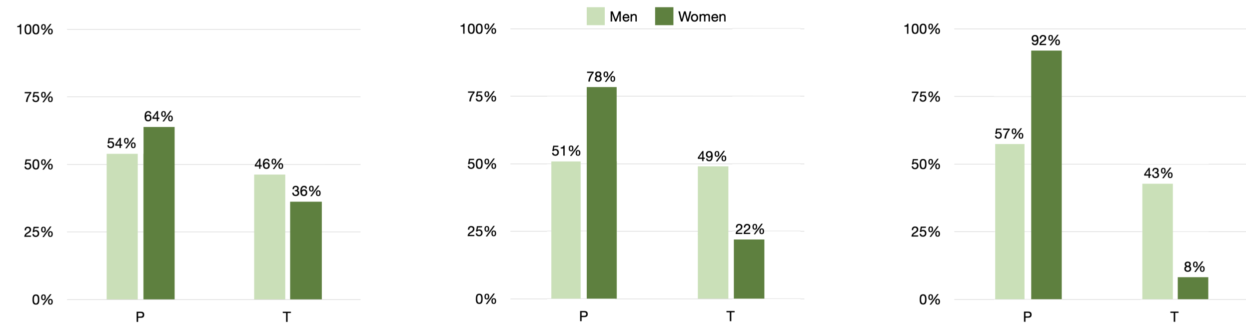
In Section 3.1 below we describe the collected data. Then, in Section 3.2, we proceed to the results of the statistical analysis using the mixed effects model for binary data described in Section 2.8.

3.1 Collected Data on Student Preference

Table 3 shows the number of assignment versions (People vs Things) handed in by men and, respectively, by women. We see that 115 men and 37 women, 152 in total, handed in assignment A (which, recall, was not part of the experiment and therefore only existed in one version). For the subsequent assignments B–D, we split the tally according to the preference for the People vs Things versions. Hence, for assignment B, 62 men selected the People version, while 53 men selected the Things version. Among the women, 23 opted for the People version of assignment B, while notably fewer, only 13, chose to do the Things version. For assignments C and D, we see a similar picture: the choices of the men distributed roughly uniformly among the two versions (albeit with slightly more men opting for the People version of assignment D), while the women consistently prefer the People over the Things versions. (For assignment C, a single student handed in both the People and Things version. This data point has been excluded from Table 3 along with those that were not handed in.)

Figure 2 shows the percentage distribution of students and indicates that the probability of choosing each People vs Things varies according to both gender and assignment. For all assignments considered, a majority of the women appear to prefer the People over the Things version of the assignments, whereas men, as a whole, do not appear to have a particular preference for any of the versions.

To summarize the variation between assignments, consider the odds of a student choosing the People version rather than the Things version. Among women choosing for assignment B, the odds are 64%/36% = 1.8, whereas for men they are 54%/46% = 1.2. So, as Figure 2 illustrates, both genders have a slight preference for the People version, as seen by their odds being greater than one. However, the women appear to have a more pronounced preference than the men, as quantified by their odds of choosing the People version being 1.8/1.2 = 1.5 times higher than the odds among men.



(a) Assignment B. **(b) Assignment C.** **(c) Assignment D.**

Figure 2: Percentage distribution of students according to their gender and the choice of version (People vs Things) for each of the three assignments.

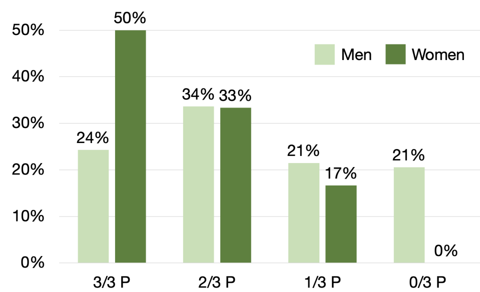


Figure 3: Overall gender preferences among the 143 students who handed in *all* three assignments. "3/3 P" abbreviates three People version choices; whereas "1/3 P", e.g., means one People choice and two Things choices.

Similarly, we find from Figure 2 that, for assignment C, the women have 3.5 times higher odds of choosing the People version than do the men, and for assignment D where only three of the 37 female students chose to complete the Things version, the odds are 8.4 times higher than those for the men.

To offer some insight into the consistency of student preference across assignments, Figure 3 shows a summary of how many times out of three each student chose the People version of an assignment. For this, we consider only the N=143 students who handed in *all* (three) assignments. We note that half (50%) of the 36 women always chose the People version, whereas the same was true for only about a quarter (24%) of the 107 men. Equally interestingly, none of the women always opted for the Things version, as compared to about a fifth (21%) of the men.

3.2 Statistical Analysis of Student Preference

For any student, given a specific assignment, B–D, we may speak about their probability of choosing the People version of that assignment over the Things version and use this as a measure of their preference. If the probability that the student chooses the People version is more than a half, we may say that they *prefer* the People version over the Things version.

A simple indication of these probabilities was given in Figure 2 in terms of empirical probabilities, from which we noted a clear preference for the People version in several groups.

Student preferences and the variation therein. Our statistical model takes into account that there may be some heterogeneity among students and that their preferences (as quantified by the probability of choosing a People version) may thus generally be at the higher or at the lower end for all three assignments. We may think of it in terms of the model capturing personal preferences as a variation of the probability of choosing the People version around some overall probability within each combination of gender and assignment.[4]

We may thus speak about an *average* or *median* student, meaning that 50% of students have a lower probability of choosing the People version. Figure 4 shows the estimated probabilities that a median student of a given gender chooses the People, respectively, the Things version of a specific assignment B–D. Consider first the *median*-student (bullets, solid line): For males (top row of Figure 4), particularly assignments B and C, the probabilities are roughly equal (horizontal solid lines), which indicates an indifference to the type of assignment, but we see that the student heterogeneity implies that at least half of the male student population actually exhibit a preference for the People version. For females (bottom row of Figure 4), however, we see a clear preference for the People version of each assignment (tilted solid lines); indeed a median female student has a higher probability of choosing People than of choosing Things in all three assignments. Note the graphical representation of each pair of probabilities in Figure 4: The steeper the slope of the line, the more pronounced the preference.

Similarly, we may consider *extreme* preferences, here taken to be 5^{th} and 95^{th} percentiles for the probability of choosing the People version (dashed: 5^{th} percentile, dotted: 95^{th} percentile). This gives an interval of probabilities (Figure 4, open circles) with the interpretation that 90% of the student population would have a personal probability within this interval; 5% of students would exhibit a lower (and more extreme) probability and another 5% of students would exhibit higher (and more extreme) probability.

[4] as implied by the assumed Gaussian variation at the log-odds scale.

The variation in student-specific preferences imply also that some women are more inclined to choose the THINGS version than the PEOPLE version (however slight this inclination may be). Indeed, using the fitted statistical model, we may readily ask how big a proportion of student have the same or smaller probability of choosing the PEOPLE version than the THINGS. For assignment B, it would be about 25% of the female student population, whereas for assignment C and D, this holds for only about 7% and less than 1%, respectively.

Gender effects in student preferences. We saw in Section 3.1 and, again, in Figure 4 that the data suggests that women are more inclined to the PEOPLE version than are the men, and also that the inclination may differ between the three assignments. A likelihood ratio test in the mixed effects model confirms that the effect of gender does vary significantly with the three assignments given ($p = 0.03$).

As before, we consider the odds of a student choosing the PEOPLE version rather than the THINGS version. Taking two students, we may then quantify how much stronger is the preference for the PEOPLE version for one student than for the other through their odds ratio. An odds ratio of about one means that their preference is of similar strength. Since student-specific preferences vary, we will naturally get some variation in the odds ratios obtained by comparing specific pairs of students, just as for the probabilities in Figure 4. Table 4 shows the median odds ratios obtained when comparing two random students, one male and one female, in their preference for a specific task. A median odds ratio of 1.7 implies that for half of such comparisons, the odds of the woman choosing the PEOPLE version of assignment B would be more than 1.7 times higher than the odds of the man choosing the PEOPLE version of assignment B.

Across all three assignments we see a clear difference in the preferences for male and female students; indeed women are more inclined to the PEOPLE version as is testified by the odds ratios all being greater than one (OR=1 corresponds to indifference in terms of equal probabilities for the two versions). For assignments C and D, the median odds ratios are 4.6 and 11.9, respectively, and this difference between genders in their preference for a PEOPLE version is statistically significant for both tasks ($p = 0.004$ and $p = 0.0005$, respectively). The gender effect for assignment B is *not* statistically significant ($p = 0.28$); indeed, we note that the confidence interval stated in Table 4 also includes values less than 1, indicating that, in principle, the median odds ratio between the female and the male student could be as low as 0.65 (i.e., the odds for the female choosing the PEOPLE version could, in principle, be only 0.65 times the odds for the male).

The variation in student-specific preference implies a considerable variation in the odds ratios quantifying the gender effect. This variation may be quantified through an interval odds ratio (IOR) [23], which with 90% probability contains the odds ratio we would find by comparing the preference of a random female student to the preference a random male student for a specific assignment. These 90% IOR are given in Table 4.

Table 4: Odds ratios quantifying the gender effect. Also shown are 95% confidence intervals (CI) indicating the uncertainty of the estimated parameters, and a 90% credible interval, interval odds ratio (IOR), giving the variability in odds ratios resulting from individual student preferences.

ID	ODDS RATIO	95% CI	p-VALUE	90% IOR
B	1.7	0.65 – 4.4	0.28	0.14 – 20
C	4.6	1.6 – 13	0.004 ⋆⋆	0.39 – 54
D	11.9	3.0 – 48	0.0005 ⋆⋆⋆	1.0 – 141

The variation in student-specific preferences also implies a within-gender variability, in that if we were to take two random students–*both male* or *both female*–and consider their preferences for a specific assignment, the student with the highest odds of choosing the PEOPLE version would have about 2.7 times higher odds than those of the other student. Naturally, this within-gender effect can be made even more pronounced by picking specifically students from each of the two extreme ends of preferences.

We note, in passing, that the probabilities indicated in Figure 4 and the 90% interval of odds ratios (IOR) in Table 4 do not take into account any variability arising from the uncertainty of the estimated model parameters; only variability arising from student heterogeneity is considered. The uncertainty about estimated parameters is reflected in the confidence intervals (CI) given in Table 4.

4 THREATS TO VALIDITY

We now consider the validity of our observational study in terms of *construct*, *internal*, and *external* validity.

4.1 Construct Validity

Establishing gender? We asked the participants to record their gender, where options included `non-binary` and `not listed` in addition to `female` and `male`. Although our study adopted a simplistic, binary perspective on gender (women xor men), we recognize that these are in no way the only existing gender identities.

Measuring preference? We quantify the preference of an individual student in terms of their choice to complete and hand in one assignment version over another; i.e., as a choice between a PEOPLE version versus an isomorphic THINGS version of an assignment. Note that, unlike in previous research [13], the subjects also had to complete the assignment which entails significantly higher commitment as opposed to the low commitment choice of simply picking a hypothetical task based on superficial appeal (something that was the case in our previous study).

4.2 Internal Validity

Preference influenced by peers? The students had the option of working on assignments on campus around their fellow classmates. This poses the risk that the choice of one student could be influenced by that of another student. To reduce this risk, we explicitly stated that each student was to pick the version they themselves found most interesting (cf. Section 2.7). The students were asked to self-report with whom, if any, they had collaborated. Collaborations

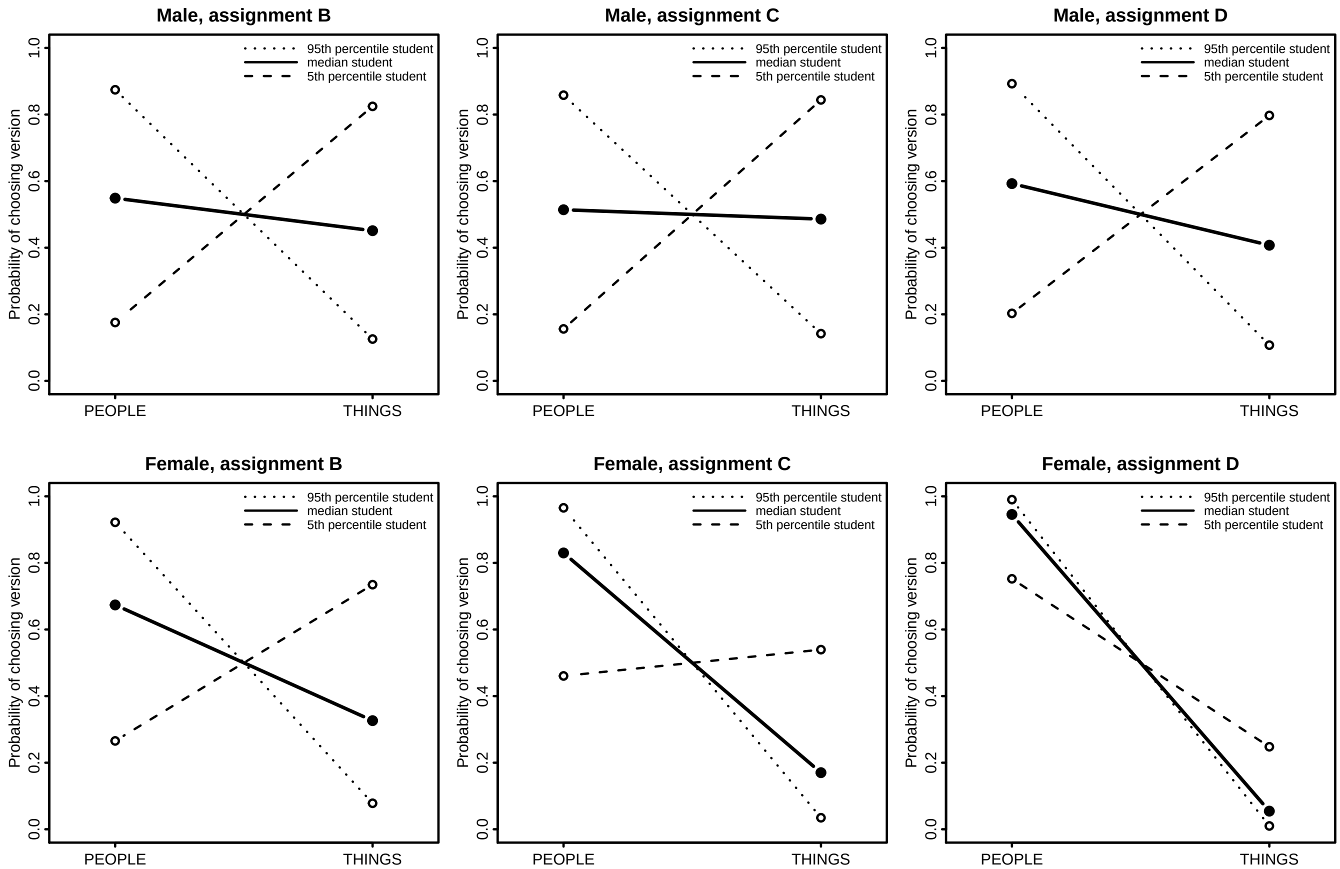


Figure 4: Variability of personal preference among students. Shown are the probabilities for median students as well as students with lowest (5th percentile) and highest (95th percentile) preference for People.

were generally diffuse, so so this information has been excluded in further analysis. In the few consistent collaborations identified, we note that the students chose different versions.

Gender questions inducing subject self-stereotyping? When asked for their gender at the beginning of an experiment, there is a documented risk that people might "self-stereotype" and answer more in accordance/disagreement with the stereotype of their gender [36]. To mitigate this threat, we asked the students for their gender *after* the experiment was over, independently of them choosing which assignment to make.

Bias from the order of presentation? The first time a student encounters the two versions (and most likely decides between them) is when opening the downloaded assignment folder. In this folder, the two versions of the assignment have presumably been presented in alphabetical order (depending on the setting of their computer's operating system). In case the student's version preference is of complete indifference, they may well just pick the first choice on the list. As mentioned in section 2.7, we ensured one version would not be alphabetically on top for all three assignments. Although some students could have simply picked the first (top) version, no clear effect from presentation order is present in the data of the students' choices.

Preference influenced by novelty or familiarity? Although participants were first year students, they had been subjected to several educational activities before receiving the first of our three assignments. Had these activities only addressed the theme of Things, for instance, there is a risk that a "novelty-effect" could have led some students to choose the People-version of our assignments, just for the sake of variation. In contrast, familiarity with the theme of Things in computing could have led other students to choose the Things-version to avoid variation. However, we do not expect this risk to play a big role, as none of the educational activities in the Introductory Programming course were themed specifically with People or Things in mind.

Subjects deliberately interfering with study outcome? If we had given the students information about the nature of the study beforehand, there is a risk that politically inclined students may have wanted to influence the study via their choices. For this reason, we did not disclose the objective of the study at any point during the experiment. When presented with the assignments, they merely knew that there were two versions of the assignment to choose from. In fact, many students enthusiastically reported that they found it interesting that they had something to choose between.

4.3 External Validity

Beyond mere preference? In this study, we did not look into whether the assignment version had an impact on the student's performance in the assignment. (Note that, in Denmark, this type of mandatory assignments are there to provide incentives for the students to work with relevant topics, in practice, during the course and are, as such, usually graded simply as "approved" vs "not approved" with the vast majority of the students getting the work approved. This does thus not provide the ideal setup for investigating performance.) It would be interesting to explore whether or not a preference for either People or Things also heightens that student's performance in the preferred version of the assignment.

Beyond first-year students? We expect to see a similar pattern for more experienced software development students. However, it may be the case that the gender preference discrepancy diminishes, as students learn to think in more abstract terms and start perceiving the problem beyond the "superficial" thematic scenario wrapping of the educational activity. We intend to investigate this, tracking the gender effect as a function of the number of years at university.

Beyond Danish students? We expect our conclusions to generalize to other Western societies with similar trade-offs between intrinsic vs extrinsic motivations influencing vocational choice.

Beyond Software Development? We hypothesize that our results could apply to not only other computing degrees, but other technical science fields, including those of STEM. However, this needs independent investigation.

5 RELATED WORK

The retention of female students in computing (and STEM in general) has been vastly explored over the last decades. Much of the literature points to factors outside of school curricula, such as satisfaction with the degree [5, 9, 20] sense of belonging [5, 6, 20], and institutional/departmental choices [14, 22]. Some studies give specific recommendations on recruitment and retention interventions for female students in computing. For example, "*take differences in experience into account*" [29], "*increase early research opportunities*" [15], and "*role models in computer science or computing for women and underrepresented groups*" [33]. However, differently from our work, most of these recommendations affect elements other than the academic content of computing programmes.

In closer relation to our work, some studies have looked at ways of increasing retention with changes to in-classroom activities and extracurricular activities [2, 10, 29, 30, 32, 35]. Some researchers and educators have looked specifically into improving retention through computing assignments: Layman et al. [24] looked at ways of making assignments more meaningful (aiding or producing value to society) to students, Bouvier et al. [8] discussed the effect of context in assignment descriptions, and Zeitz and Anewalt [39] looked at designing assignments to increase students' awareness of diversity, inclusion, and accessibility.

Although not related to gender, it is also worth mentioning Lee and Ko's study on engagement in a programming game [25]. They found that using vertebrate and invertebrate game elements (animals), in contrast to inanimate elements (things), kept players engaged for longer.

Similarly to ours, a few other studies have investigated preference for certain themes in computing assignments. Wilson, B. C. [38] found that female students prefer real-life application-related assignments over game-related assignments. Rader et al. [34] found that students generally preferred "fun" assignments over those involving socially relevant or practical elements and that, although female students had a stronger interest in the socially relevant projects than male students, other projects still had higher appeal. In contrast to these studies, our study asks the students to compare two versions of the same assignment, differing only in theme, but not in the computing problem itself. This eliminates some of the factors that could influence student preferences when picking one assignment over another.

Designing curricula that relate to the "real world" and have an impact on society is mentioned in several studies [24, 29, 38]. It is important to note, however, that the People perspective is not the same as relating to society. For instance, an assignment on *"the shortest path to lay down power lines between cities"* would be society-, but not People-related; in contrast, an assignment on *"the fastest way for a person to walk to every cinema in the country"* would be People-, but not society-related. An assignment that relates to both could be *"the best public-transport route to avoid overcrowding."* In short, our work adds quantitative data to the archive of research carried out before ours. It should be seen as complementary to the above mentioned strategies in increasing retention among women in computing.

To the best of our knowledge, this study is the first to explore retention through gender differences in preference for People or Things in computing assignments.

6 CONCLUSION

In summary, there are three main highlights from our analysis:

> **Observation 1:** Women are more inclined to chose the People version than men are.
>
> **Observation 2:** The strength of preferences depends on the task at hand. In particular, the effect of gender (the odds ratio) observed in our study varies from 1.7 to 11.9 on the three assignments given.
>
> **Observation 3:** There is considerable variation between students.

Further reflection on these observations lead to a number of additional points that are worth mentioning as they can serve as inspiration for further study. First of all, further research should look into whether the heterogeneity between students can be explained by factors not measured by this study. For instance, our previous study [13] indicated that self-perceived programming experience has an effect on the preference, and it is very possible that such effects can also be measured in freshers at university (e.g., the progression/experiences through the semester may well differ for the two genders). This is important as it is possible that the students who are more experienced became "conditioned" by the type of materials they were exposed to (during prior exercises or as a part of self-education). Second of all, the difference we see between assignments may be due to characteristics of the assignments or it could very well be due to some underlying temporal effect (such as the difference in progression/experiences). At least some thoughts on why the three assignments differ in regards to preferences are needed. One possibility, which could be further studied, is that assignment D uses a more narrative style and that this may have influenced the students. The reason for it is that narrations are typically associated with characters (and not with objects). In addition, assignment D is also the only task that presents a holistic scenario of creating a whole suite of functions completely from scratch.

Third of all, it would be interesting to learn more about the extreme subsets of students (the upper and lower 5th percentiles). It is especially interesting to see if we can pinpoint some relevant similarities between them (for example their programming experience level). For diversity, it is important to assess how much the change of education strategy we suggest affects these "extremity groups."

As our study is observational, the results may not readily generalize to a larger population than contemporary students at the IT University of Copenhagen. However, our findings are remarkably consistent, not only with our previous study in the context of recruitment concerning high-school students [13], but also with Lippa's more general studies involving a twin study from 1976 and American psychology students from the 1990s [27] and a later study from 2010 [28].

The overarching point is that themes matter, although the specific thematic instantiation naturally affect the magnitude of the appeal. We recommend:

> **<u>Recommendation</u>:** Educators should consider favoring themes involving People over those involving Things in computing educational activities, as it appears to increase the appeal of the activities to women.

We expect this generalizes to all teaching/learning activities, such as assignments, exercises, examples, projects, demonstrations, analogies, and illustrations.

Our study uncovered a notable individual difference in the preferences for People vs Things tasks. Individual women may prefer Things tasks just like individual men might may prefer People tasks. To accommodate this variation, we suggest providing students with *both* types of assignments; i.e., People *and* Things. Different students will thus be able to make different choices. We hope that diversifying the educational activities could ultimately help diversifying the student body involved with computing.

ACKNOWLEDGMENTS

We thank the students attending in the Introductory Programming course at the IT University of Copenhagen during the fall of 2020. We also thank the Teaching Assistants of the course—Emil Joakim Jensen Bartholdy, Jakob Thorsen Staugaard, Laura Caroline Cholvat, Nynne Grauslund Kristiansen, and Sebastian Mateos Nicolajsen—for their support and collaboration.

REFERENCES

[1] Alan Agresti. 2013. *Categorical data analysis* (3rd ed.). John Wiley & Sons, USA.

[2] Lecia J. Barker, Melissa O'Neill, and Nida Kazim. 2014. Framing classroom climate for student learning and retention in computer science. In *Proceedings of the 45th ACM technical symposium on Computer science education - SIGCSE '14*. ACM Press, New York, New York, USA, 319–324. https://doi.org/10.1145/2538862.2538959

[3] Camilla Persson Benbow, David Lubinski, Daniel L Shea, and Hossain Eftekhari-Sanjani. 2000. Sex differences in mathematical reasoning ability at age 13: Their status 20 years later. *Psychological science* 11, 6 (2000), 474–480.

[4] Sylvia Beyer. 2014. Why are women underrepresented in Computer Science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. *Computer Science Education* 24, 2-3 (2014), 153–192.

[5] Maureen Biggers, Anne Brauer, and Tuba Yilmaz. 2008. Student perceptions of computer science: a retention study comparing graduating seniors with cs leavers. *ACM SIGCSE Bulletin* 40 (2 2008), 402–406. Issue 1. https://doi.org/10.1145/1352322.1352274

[6] Jacob Clark Blickenstaff*. 2005. Women and science careers: leaky pipeline or gender filter? *Gender and Education* 17 (10 2005), 369–386. Issue 4. https://doi.org/10.1080/09540250500145072

[7] BlueJ. 2020. About BlueJ. https://www.bluej.org/about.html

[8] Dennis Bouvier, Ellie Lovellette, John Matta, Bedour Alshaigy, Brett A. Becker, Michelle Craig, Jana Jackova, Robert Mccartney, Kate Sanders, and Mark Zarb. 2016. Novice programmers & the problem description effect. In *Proceedings of the 2016 ITiCSE Working Group Reports, ITiCSE 2016*. 103–118. https://doi.org/10.1145/3024906.3024912

[9] Eileen D. Bunderson and Mary Elizabeth Christensen. 1995. An Analysis of Retention Problems for Female Students in University Computer Science Programs. *Journal of Research on Computing in Education* 28 (9 1995), 1–18. Issue 1. https://doi.org/10.1080/08886504.1995.10782148

[10] Jeffrey C. Carver, Lisa Henderson, Lulu He, Julia Hodges, and Donna Reese. 2007. Increased Retention of Early Computer Science and Software Engineering Students Using Pair Programming. In *20th Conference on Software Engineering Education & Training (CSEET'07)*. IEEE, 115–122. https://doi.org/10.1109/CSEET.2007.29

[11] Sapna Cheryan. 2012. Understanding the paradox in math-related fields: Why do some gender gaps remain while others do not? *Sex Roles* 66, 3-4 (2012), 184–190.

[12] Sapna Cheryan, Sianna A Ziegler, Amanda K Montoya, and Lily Jiang. 2017. Why are some STEM fields more gender balanced than others? *Psychological bulletin* 143, 1 (2017), 1.

[13] Ingrid Maria Christensen, Melissa Høegh Marcher, Paweł Grabarczyk, Therese Graversen, and Claus Brabrand. 2021. Computing Educational Activities Involving PEOPLE Rather Than THINGS Appeal More to Women (Recruitment Perspective). In *Proceedings of the 17th ACM Conference on International Computing Education Research (ICER 2021), August 16–19, 2021, Virtual Event, USA*. ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3446871.3469758

[14] J. McGrath Cohoon. 1999. Departmental differences can point the way to improving female retention in computer science. In *The proceedings of the thirtieth SIGCSE technical symposium on Computer science education - SIGCSE '99* (New York, New York, USA). ACM Press, 198–202. https://doi.org/10.1145/299649.299753

[15] Tanya L. Crenshaw, Erin Wolf Chambers, and Heather Metcalf. 2008. A case study of retention practices at the University of Illinois at Urbana-Champaign. In *Proceedings of the 39th SIGCSE technical symposium on Computer science education - SIGCSE '08* (New York, New York, USA). ACM Press, 412–416. https://doi.org/10.1145/1352135.1352276

[16] Amanda B Diekman, Elizabeth R Brown, Amanda M Johnston, and Emily K Clark. 2010. Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological science* 21, 8 (2010), 1051–1057.

[17] Informatics Europe. 2019. Informatics Bachelor's degrees awarded to female students - Percentage. https://www.informatics-europe.org/data/higher-education/statistics/bachelor_degrees_awarded_percentage.html

[18] Eurostat. 2018. ICT specialists: statistics on hard-to-fill vacancies in enterprises - Statistics Expained. https://ec.europa.eu/eurostat/statistics-explained/index.php/ICT_specialists_-_statistics_on_hard-to-fill_vacancies_in_enterprises#Employment_and_recruitment_of_ICT_specialists

[19] Eurostat. 2019. ICT specialists in employment. https://ec.europa.eu/eurostat/statistics-explained/index.php/ICT_specialists_in_employment#ICT_specialists_by_sex

[20] Michail N. Giannakos, Ilias O. Pappas, Letizia Jaccheri, and Demetrios G. Sampson. 2017. Understanding student retention in computer science education: The role of environment, gains, barriers and usefulness. *Education and Information Technologies* 22 (9 2017), 2365–2382. Issue 5. https://doi.org/10.1007/s10639-016-9538-1

[21] Diane F Halpern, Camilla P Benbow, David C Geary, Ruben C Gur, Janet Shibley Hyde, and Morton Ann Gernsbacher. 2007. The science of sex differences in science and mathematics. *Psychological science in the public interest* 8, 1 (2007), 1–51.

[22] Kulli Kori, Margus Pedaste, Eno Tonisson, Tauno Palts, Heilo Altin, Ramon Rantsus, Raivo Sell, Kristina Murtazin, and Tiia Ruutmann. 2015. First-year dropout in ICT studies. In *2015 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 437–445. https://doi.org/10.1109/EDUCON.2015.7096008

[23] Klaus Larsen, Jørgen Holm Petersen, Esben Budtz-Jørgensen, and Lars Endahl. 2000. Interpreting Parameters in the Logistic Regression Model with Random Effects. *Biometrics* 56 (2000), 909–914.

[24] Lucas Layman, Laurie Williams, and Kelli Slaten. 2007. Note to self: make assignments meaningful. In *Proceedinds of the 38th SIGCSE technical symposium on Computer science education - SIGCSE '07* (New York, New York, USA). ACM Press, 459–463. https://doi.org/10.1145/1227310.1227466

[25] Michael J. Lee and Amy J. Ko. 2012. Investigating the role of purposeful goals on novices' engagement in a programming game. In *Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC*. 163–166. https://doi.org/10.1109/VLHCC.2012.6344507

[26] Ole Lehrmann Madsen, Birger Møller-Pedersen, and Kristen Nygaard. 1993. *Object-Oriented Programming in the BETA Programming Language*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA. Chapter 18: Conceptual Framework.

[27] Richard A. Lippa. 1998. Gender-related individual differences and the structure of vocational interests: The importance of the people-things dimension. *Journal of Personality and Social Psychology* 74, 4 (1998), 996–1009. https://doi.org/10.1037/0022-3514.74.4.996

[28] Richard A. Lippa. 2010. Gender Differences in Personality and Interests: When, Where, and Why? *Social and Personality Psychology Compass* 4, 11 (2010), 1098–1110. https://doi.org/10.1111/j.1751-9004.2010.00320.x

[29] Jane Margolis and Allan Fisher. 2002. Unlocking the clubhouse: The Carnegie Mellon experience. *ACM SIGCSE Bulletin* 34 (6 2002), 79–83. Issue 2. https://doi.org/10.1145/543812.543836

[30] Briana B. Morrison and Jon A. Preston. 2009. Engagement: gaming through the curriculum. *ACM SIGCSE Bulletin* 41 (3 2009), 342–346. Issue 1. https://doi.org/10.1145/1539024.1508990

[31] K.-Y. Liang P. J. Diggle, P. Heagarty and S. L. Zeger. 2002. *Analysis of Longitudinal Data* (2nd ed.). Oxford University Press.

[32] Joan Peckham, Peter Stephenson, Jean-Yves Hervé, Ron Hutt, and Miguel Encarnação. 2007. Increasing student retention in computer science through research programs for undergraduates. In *Proceedings of the 38th SIGCSE technical symposium on Computer science education - SIGCSE '07*. ACM Press, New York, NY, USA, 124–128. https://doi.org/10.1145/1227310.1227354

[33] Markeya S. Peteranetz, Abraham E. Flanigan, Duane F. Shell, and Leen-Kiat Soh. 2018. Future-Oriented Motivation and Retention in Computer Science. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (New York, NY, USA). ACM, 350–355. https://doi.org/10.1145/3159450.3159513

[34] Cyndi Rader, Doug Hakkarinen, Barbara M. Moskal, and Keith Hellman. 2011. Exploring the appeal of socially relevant computing: Are students interested in socially relevant problems?. In *SIGCSE'11 - Proceedings of the 42nd ACM Technical Symposium on Computer Science Education*. 423–428. https://doi.org/10.1145/1953163.1953288

[35] Eric S. Roberts, Marina Kassianidou, and Lilly Irani. 2002. Encouraging women in computer science. *ACM SIGCSE Bulletin* 34 (6 2002), 84–88. Issue 2. https://doi.org/10.1145/543812.543837

[36] Stacey Sinclair, Curtis D Hardin, and Brian S Lowery. 2006. Self-stereotyping in the context of multiple social identities. *Journal of Personality and Social Psychology* 90 (2006), 529. Issue 4. https://doi.org/10.1037/0022-3514.90.4.529

[37] Fay Pedersen Tveranger. 2017. *Programming in school. An insight to the Norwegian programming pilot and the inclucion/exclusion of girls in computer programming education.* Master's thesis. The University of Bergen.

[38] Brenda Cantwell Wilson. 2006. Gender Differences in Types of Assignments Preferred: Implications for Computer Science Instruction. *Journal of Educational Computing Research* 34 (4 2006), 245–255. Issue 3. https://doi.org/10.2190/7FLU-VKJL-86RM-5RQG

[39] Jessica Zeitz and Karen Anewalt. 2021. Creating a Repository of Diversity and Inclusion Assignments for Computer Science. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (New York, NY, USA). ACM, 1321–1321. https://doi.org/10.1145/3408877.3439654

EXCERPTS FROM ASSIGNMENT B

See Figure 5.

EXCERPTS FROM ASSIGNMENT C

See Figure 6.

EXCERPTS FROM ASSIGNMENT D

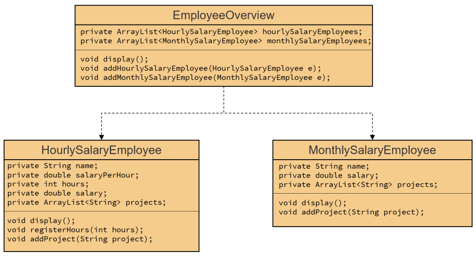
See Figure 7.

Mandatory Assignment B
Employee System

[...]

Task 1

The `EmployeeSystem` project has three classes, as can be seen in Figure 1: `HourlySalaryEmployee`, `MonthlySalary-Employee`, and `EmployeeOverview`. The classes represent a system for a company that consults large organizations in teamwork. There are two employment conditions: *hourly salary*, where the salary is calculated per working hour, and *monthly salary*, where the salary is fixed. `EmployeeOverview` provides an overview of all employment conditions.



[...]

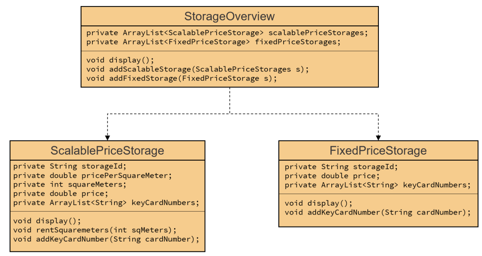
(a) People

Mandatory Assignment B
Storage System

[...]

Task 1

The `StorageSystem` project has three classes, as can be seen in Figure 1: `ScalablePriceStorage`, `FixedPriceStorage` and `StorageOverview`. The classes represent a system for a company that rents out large storage space for business use. There are two rental options: *scalable price*, where the price is calculated per rented square meter, and *fixed price*, where the price is fixed. `StorageOverview` provides an overview of all rental options.



[...]

(b) Things

Figure 5: Excerpts from assignment B in two versions

Mandatory Assignment C
Hotel Booking

[...]

Task 1

The `HotelBooking` project has four classes: `Hotel`, `Room`, `Booking` and `Demo`. The classes are part of a program that can be used in a hotel. However, the program is still in an early stage of development, so in the next four tasks you will make it less error-prone, as well as expand the program with new functionalities.

[...]

Task 4

A hotel employee wants to have a visual overview of the hotel's bookings. Create a method `void showHotel()` in the class `Hotel`, that prints a matrix representing all rooms in the hotel. If a room is booked, print "`[X]`" and if the room is available, print "`[]`". Figure 1 is a printout of a hotel with four floors and three rooms per floor, where the rooms `1.2; 2.2; 3.0; 3.1` are booked (the same as in `Demo`).

```
[X][X][ ]
[ ][ ][X]
[ ][ ][X]
[ ][ ][ ]
```

[...]

(a) People

Mandatory Assignment C
Cargo Ship Booking

[...]

Task 1

The `CargoShipBooking` project has four classes: `CargoShip`, `Container`, `Booking` and `Demo`. The classes are part of a program that can be used on a cargo ship. However, the program is still in an early stage of development, so in the next four tasks you will make it less error-prone, as well as expand the program with new functionalities.

[...]

Task 4

A screen needs to display a visual overview of the cargo ship's bookings. Create a method `void showCargoShip()` in the class `CargoShip`, which prints a matrix representing all the containers on the cargo ship. If a container is booked, print "`[X]`" and if the container is vacant, print "`[]`". Figure 1 is a printout of a cargo ship with four rows and three containers per row, where the containers `1.2; 2.2; 3.0; 3.1` are booked (the same as in `Demo`).

```
[X][X][ ]
[ ][ ][X]
[ ][ ][X]
[ ][ ][ ]
```

[...]

(b) Things

Figure 6: Excerpts from assignment C in two versions

Mandatory Assignment D
Customer Tracker

[...]

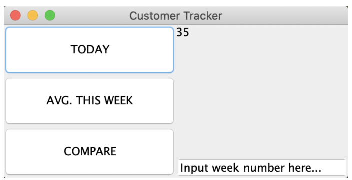
Task 1

In recent weeks, a store manager has collected data on how many customers visit the store in the course of a day. The store manager now needs a program that is able to process and display the data collected.

[...]

Task 4 (Optional)

The system now needs to be expanded with a user interface *(Graphical User Interface (GUI))* so that a user can interact with the program.



[...]

(a) People

Mandatory Assignment D
Electricity Tracker

[...]

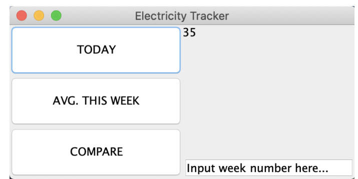
Task 1

In recent weeks, an electricity meter has collected data on how much electricity is used in the building in the course of a day. The electricity meter now needs a program that is able to process and display the data collected.

[...]

Task 4 (Optional)

The system now needs to be expanded with a user interface *(Graphical User Interface (GUI))* so that a user can interact with the program.



[...]

(b) Things

Figure 7: Excerpts from assignment D in two versions