

Article

Strengthening Science Understanding with Learning Trails

Wolfgang Leister¹ , Ingvar Tjøstheim¹, Göran Joryd^{2,3}, Jan Alfred Andersson⁴, and Håvard Heggelund⁴

¹ Norsk Regnesentral, Postboks 114 Blindern, 0314 Oslo, Norway; wolfgang.leister@nr.no (W.L.); ingvar.tjostheim@nr.no (I.T.);

² Expology AS, Sagveien 23F, 0459 Oslo, Norway; goran@expology.no (G.J.)

³ Museum of Cultural History, Postboks 6762 St. Olavs plass, 0130 Oslo, Norway; goran.joryd@khm.uio.no (G.J.)

⁴ Norwegian Museum of Science and Technology, Kjelsåsveien 143, 0491 Oslo, Norway; Jan.Andersson@tekniskmuseum.no (J.A.A.); Haavard.Heggelund@tekniskmuseum.no (H.H.).

* Correspondence: wolfgang.leister@nr.no (W.L.)

Version May 15, 2019 submitted to Preprints

Abstract: The Norwegian Museum of Science and Technology have developed a learning concept for school classes in science centres named 'learning trails'. In this concept, groups of students perform a series of thematically related experiments with installations in the science centre. The learning trails are designed to support the generic learning outcomes for science centre visits. We argue for using the previously developed Engagement Profile in an indicator to determine both media forms and generic learning outcomes for such learning concepts. Further, we implemented the learning trails in two modes: one mode used paper-based content to guide the students, while the other mode supported the use of tablet PCs where engaging content is triggered when the students approach the location of an experiment in the learning trail. We studied the engagement factors of the learning trails and observed how school classes use these. In a study with 113 students from lower secondary school, they answered short questionnaires that were integrated into the implementation of the learning trails. While the concept of the learning trails was evaluated positively, we could not find significant differences in how engaging the two implemented modes were.

Keywords: learning trail; science centres; visitor engagement; generic learning outcomes

1. Introduction

Science centres are informal learning environments [1] that offer exhibits supporting free-choice learning, as well as specific programmes for organised school class visits. We want to explore to what degree such learning programmes engage students and find means to strengthen the engaging factors of an exhibit [2].

The Norwegian Museum of Science and Technology (NTM) introduced the concept of learning trails around science centre installations grouped thematically. The learning trails combine physics experiments with technology history from the exhibitions in the science centre. The motivation for this is to foster learning from using these installations and to create dialogues and narratives [3] that explain science phenomena. The activities of the learning trails are designed for groups of up to four students at a time.

The objective of this paper is to explore how the elements in the learning trail concept can be used to create engaging content that supports the generic learning outcomes [4] specified by the science centre. We relate the Engagement Profile [2] to the generic learning outcomes and evaluate the learning trails in an empiric study with students from secondary school classes. Further, we investigate the impact of presenting location-based content on a tablet PC versus a paper-based version of the learning trails. To study this, we implemented a prototype that presents content upon arrival at an experiment, using unobtrusive and affordable in-door location technology.

33 In the following, we present an overview of related work including a review of learning outcomes,
34 the Engagement Profile, and location technology in museums and science centres (Section 2), before we
35 show how to translate the Engagement Profile to terms related to media forms and learning outcomes
36 (Section 3). Thereafter, we present the concept of the learning trails and their implementation (Section 4)
37 and develop the Engagement Profiles for the learning trail concept and the single experiments, before
38 deriving the related media forms (Section 5). Section 6 presents an empiric study where the prototype
39 of the learning trails was evaluated with students from secondary school classes. Section 7 concludes
40 our essay.

41 2. Related Work

42 We focus on school children visiting science centres as a class activity as the main target group for
43 our work. As these students may have diverse learning agendas and prefer diverse ways of learning
44 (such as reading, interacting with others, touching and doing) [4], science centres offer a diversity of
45 exhibits that can be explored during the visit. The design of installations [5] should address factors
46 such as the targeted learning outcome, learning styles, levels of engagement, and context.

47 2.1. Learning Outcomes in Science Centres

48 Learning outcomes in science centres are difficult to specify and to measure, even if we consider
49 the compulsory participation of school children. The pedagogical tool Generic Learning Outcomes
50 (GLO) [4] describes the impact of learning in museums in terms of a) knowledge and understanding,
51 b) skills, c) change in attitudes and values, d) enjoyment, inspiration and creativity, and e) activity,
52 behaviour and progression. These five aspects cannot be used as guidelines to control the design
53 process of exhibits, as these aspects are yet unrelated to design properties. However, we note that
54 design and implementation elements will have a considerable impact on the GLO. Therefore, we seek
55 a way to predict how changes in design and implementation will impact the GLO.

56 Brown [6] remarked that the GLO are subjective, do not measure learning directly, and are
57 most effective as post hoc measures, that is, after the learning experience. He suggested to consult
58 Laurillard's taxonomy of educational media [7] that is based on the teacher's concepts and constructed
59 environment, and the student's concepts and specific actions. The four kinds of activities in her
60 framework are: discussion, adaptation, interaction, and reflections. Unlike the GLO, Laurillard's
61 framework is not specifically developed for science centres, and adjustments for its use in science
62 centre learning might be necessary.

63 In Laurillard's framework, the related learning experiences and adjacent media forms (in
64 parentheses) are: i) attending, apprehending (narrative); ii) investigating, exploring (interactive);
65 iii) discussing, debating (communicative); iv) experimenting, practising (adaptive); and v) articulating,
66 expressing (productive).

67 As all five learning experience types and media forms are present in science centre learning,
68 Laurillard's framework can be applied to exhibits in science centres. However, these learning
69 experiences are present to a varying degree in each exhibit or concept. Experimenting and practising,
70 as well as investigating and exploring are the most prevalent forms of experiences, while specific
71 concepts, such as the learning trails, can extend learning to other learning experience types. Notice
72 that the work by Laurillard [7] and that by Brown [6] were published before many of the current media
73 technologies were introduced in science centres; thus, the methods and technologies described there
74 might be somewhat outdated and could need adjustment.

75 The GLO are used as a basis for evaluations in science centres and museums. For instance,
76 Ayudhya and Vavoula [8] use the GLO to guide the design of questions about the outcomes in an
77 assessment of a mobile app used by families in a science museum. In their analysis, they also encode
78 observations captured on video according to Bitgood [9]'s attention-value model, that comprises of the
79 actions: capture, focus, and engage (and an extra class: engage together [8]). Visitor observations and

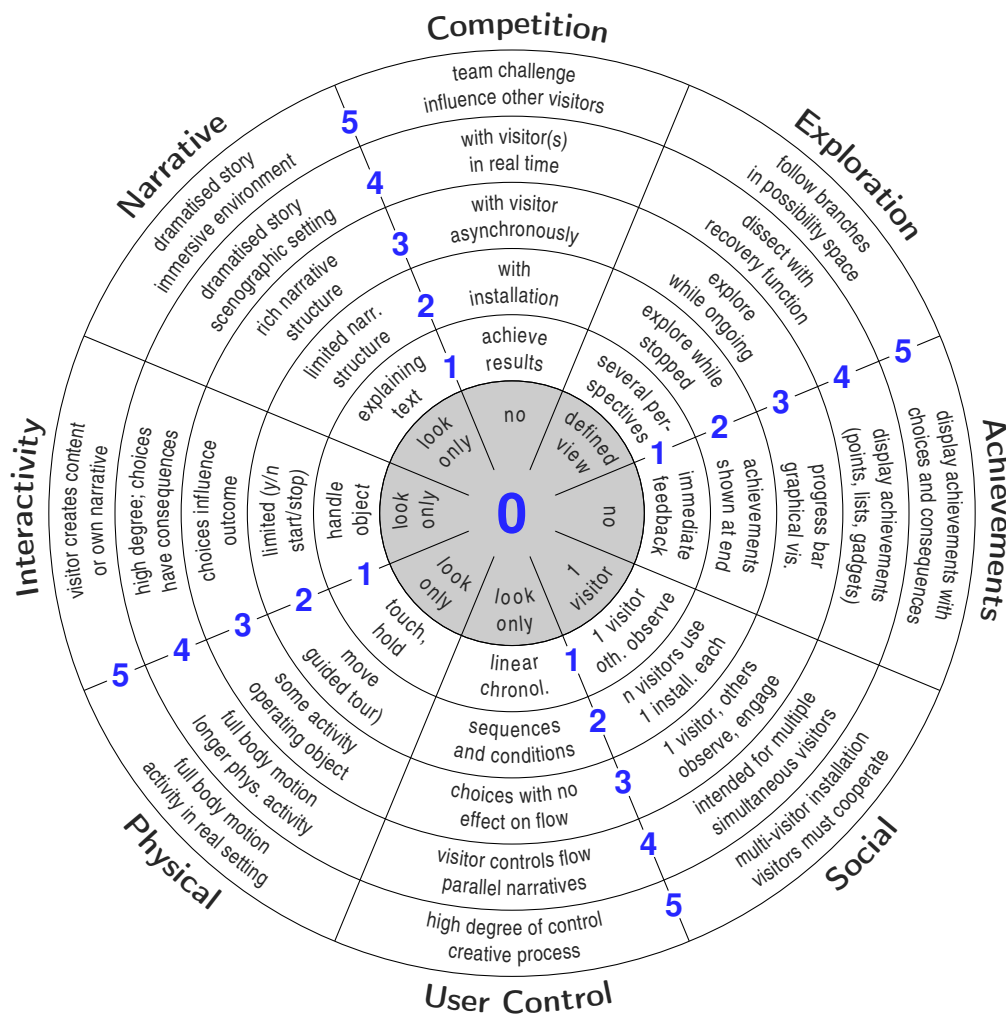


Figure 1. The dimensions of the Engagement Profile explained with short definitions. To define the value of a property, find the adjacent number of the phrases that fit best.

80 encoding activities belong to the visitor-centric view of assessment. In contrast, we want to focus on
 81 the installation-centric view that is discussed by Leister *et al.* [2, p.51]).

82 2.2. Quantitative Evaluation of User Engagement

83 Behavioural engagement is one of the factors that has a positive correlation to achievement-related
 84 outcomes [cf. 10, p. 70ff]. In informal learning arenas, this implies that engaging exhibits and
 85 installations will foster better learning outcomes than exhibits that do not engage.

86 To evaluate how engaging installations are, the Engagement Profile [2] has been used alongside
 87 sensors, observations, and questionnaires to measure engagement and user satisfaction. The
 88 Engagement Profile has been applied to the design process at museums and science centres [5], to
 89 analyse engagement and narratives for installations [3], and to develop a robotic teaching assistant for
 90 students at a university college [11].

91 The Engagement Profile quantifies the characteristics of installations along eight dimensions, each
 92 of which is given a value between 0 and 5. The dimensions of the Engagement Profile represent the
 93 degrees of *competition* (C), *narrative* elements (N), *interaction* (I), *physical* activity (P), *user control* (U),
 94 *social* aspects (S), *achievements* awareness (A), and *exploration* possibilities (E). A graphical presentation
 95 of the Engagement Profile is shown in Figure 1 as a reference.

96 External influences are not taken into account in the Engagement Profile since these are not
97 properties of the direct learning environment. Physical factors, such as noise, light or smell could
98 play a role in the perception of engagement, but need to be handled outside the Engagement Profile.
99 Properties that belong to the context, such as social factors, institutional factors, or recent incidents
100 personally or globally, are excluded. However, these factors still need to be taken into account in the
101 assessment process, e.g., as suggested in the DEX-framework [12].

102 2.3. Location-Based Systems in Science Centres and Museums

103 We posit that engagement can be increased when offering tailored content based on the visitor's
104 current location. Such location-based systems are commonly used as a component of installations, as
105 well as part of visitor studies and exhibit evaluations [13]. Several authors [e.g., 14,15] suggest to adapt
106 content to the visitor's current situation or to adjust the visiting path using on-line tracking. In our
107 approach, we use on-line tracking to trigger engaging content based on the location of the visitor.

108 Location-based systems can be used to retrieve viewing times, itinerary, speed, group behaviour,
109 and so on. Yalowitz and Bronnenkant [16] presented methods for visitor tracking and timing in
110 museums. They classified the variables to be recorded into a) stopping behaviour; b) other behaviour;
111 c) observable demographic variables; and d) situational variables. Further, they addressed practical
112 issues, technology, and ethics to collect, analyse, and interpret these variables.

113 Baldwin and Kuriakose [14] presented several technologies for tracking visitors. Tracking data are
114 used to a) predict a visitor's future path in the museum; b) recommend exhibits of potential interest to
115 the visitor, and c) personalise content delivered to visitors. They also explored the impact of physical
116 proximity and visitor gaze on exhibit engagement.

117 Yoshimura *et al.* [17] presented a study where they use Bluetooth proximity data of visitors' smart
118 phones to measure the visitors' transition between places in a museum. Moussouri and Roussos
119 [18] discussed cultural itineraries of visitors and present a study where outdoor tracking devices
120 are used to extract the paths of visitors in the London zoo. Further, Moussouri and Roussos [19]
121 proposed a methodology for representing location-based data collected by the use of smart-phones.
122 They presented three ways: a) trail-based representation; b) functional representation; and c) statistical
123 distributions of displacement.

124 The prediction of visitor's sentiments and future behaviour can be based on current observations.
125 Parsons *et al.* [15] suggested to use viewing times as an indicator of preference, and they propose a
126 recommendation system based on this idea. Bohnert and Zukerman [20] used viewing times as an
127 indicator for interest. They proposed non-intrusive personalisation of the museum experience based
128 on viewing times of previous visitor behaviour and evaluated two prediction approaches.

129 Besides outlining exhibit design approaches and strategies, Bitgood [21] presented three types of
130 visitor measures of success: a) behaviour measures including stopping (attracting power), viewing
131 time (holding power), social impact, human factors impact, and trace or decay measures; b) knowledge
132 acquisition (memory, comprehension); and c) affective measures (attitude change, interest level,
133 satisfaction).

134 2.4. Indoor-location Technologies

135 In-door location can be used both for analysis, and to adapt and control the stream of
136 content to the visitor. Mautz [22] presented a variety of indoor-tracking systems based on ample
137 technologies. Lymberopoulos *et al.* [23] compared indoor-location technologies that are based on WiFi,
138 geo-magnetic, sound signals for the infrastructure-free technologies, and RF-beacons, RFID, infrared,
139 ultrasound, Bluetooth, short-range FM transmitters, lights, and magnetic signal modulators for the
140 infrastructure-based technologies. Bickersteth and Ainsley [24] compared ample technologies for
141 tracking using mobile phones in museums, such as use of the Temporary Mobile Subscriber Identity
142 (TMSI), Bluetooth, and WiFi. In museums and science centres, we have also seen location approaches
143 based on QR-codes and camera-based tracking [25].

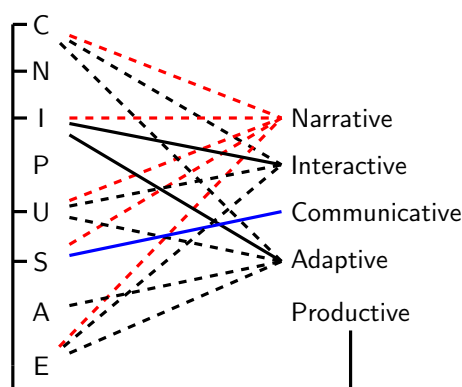


Figure 2. Aligning the entities of the Engagement Profile (left) with those of Laurillard's taxonomy (right). The dashed lines indicate partial impact. Colours are used for graphical clarity.

144 3. Generic Learning Outcomes, Media Forms, and the Engagement Profile

145 Learning outcomes depend on how students use the installations within the context of their
 146 science centre visit and how the installations and activities are designed. We have already outlined
 147 that the GLO define the outcome of a science centre visit, while Laurillard's taxonomy qualitatively
 148 describes activities and media forms for learning. In contrast, the Engagement Profile describes
 149 properties of exhibits quantitatively.

150 The implementation of an exhibit, characterised by the Engagement Profile, supports the learning
 151 outcome and the GLO through the learning activities in the science centre. Further, we posit that the
 152 media form can be described by the Engagement Profile. To support this claim, we relate the terms of
 153 Laurillard's taxonomy to the terms used in the Engagement Profile and define thresholds that indicate
 154 which media forms an exhibit has.

155 To create this indicator, we use the description of the media forms, adapted from the work by
 156 Brown [6] as a starting point. Further, we use the description of the Engagement Profile (see Figure 1
 157 for the graphical short form). We create the indicator by evaluating which term in the Engagement
 158 Profile fits best to the description of the respective media form.

159 This results in the following findings: The variables C, I, U, and E have an impact on on the
 160 narrative, interactive, and adaptive media forms. We also observed that the productive media form
 161 requires several high values in the Engagement Profile. Further, P appears to be irrelevant to indicate
 162 the media form as defined in Laurillard's framework. The relationship between the Engagement
 163 Profile and Laurillard's framework is graphically shown in Figure 2.

164 By identifying which values of the Engagement Profile suit the description of the media forms,
 165 we are able to set up conditions for which media form a given exhibit potentially can have. These
 166 conditions and their respective thresholds are shown in Table 1. The media forms relate to the
 167 Engagement Profile as follows:

168 **Narrative** media forms are described to be linear, highly structured, and non-interactive. Interestingly,
 169 the entire range of N applies for this media form. To fit into the narrative media form, most
 170 values of C, I, U, S, and E need to be rather low, i.e., between 0 and 2. Values above the threshold
 171 of 2 need to be considered regarding their impact from case to case.

172 **Interactive** media forms allow learners to explore in a non-linear way, but the content remains
 173 unchanged. This description aligns with high values of I, E and U; however, values of 5 for these
 174 three variables are not suitable. Further, medium high or high values of S and E are suitable.

175 **Communicative** media forms support feedback and foster discussions, which indicates high values
 176 of S. The other variables do not have an impact here.

177 **Adaptive** media forms adapt responses to the student's actions. This is supported by high values of I,
 178 E, and A, and medium high or high values of U and C.

Table 1. Engagement Profile indicators for media types

Laurillard's taxonomy activity	media form	Engagement Profile indicator
attending apprehending	narrative	$0 \leq z \leq 2$, for most $z \in \{C, I, U, S, E\}$; (N not significant)
investigating exploring	interactive	$3 \leq z \leq 4$, for most $z \in \{I, E, U\}$
discussing debating	communicative	$3 \leq S \leq 5$
experimenting practising	adaptive	$3 \leq y \leq 5$; $2 \leq z \leq 5$, for most $y \in \{I, E, A\}$, $z \in \{U, C\}$
articulating expressing	productive	$4 \leq z \leq 5$, for several $z \in \{C, N, I, U, S\}$

179 **Productive** media forms are tools where learners can express themselves to demonstrate their
180 understanding. This implies high values of C, N, I, U, and S.

181 In practice, supported media forms for a given exhibit can be estimated by first creating the
182 Engagement Profile, that is finding the most suitable description in Figure 1 or by using the textual
183 description in [2, p.56, Table II]. Thereafter, Table 1 is used to check whether the condition for each
184 media form is met.

185 4. Learning Trails

186 The learning trails at NTM are designed for school classes. They combine physics experiments
187 with technology history from the exhibitions, grouped thematically. To increase the learning effect,
188 these installations are intended to create a wider dialogue and narratives that explain science
189 phenomena. From the perspective of the GLO, increased knowledge and understanding are most
190 important outcomes, while the visit at the science centre shall be enjoyable and inspiring.

191 NTM has organised learning objectives for subjects that have been discussed in class before the
192 museum visit. Further, the museum also expected that students understand the task better when they,
193 additionally, can listen to content from an audio file.

194 There are indications from earlier observations at NTM that the students will be more quiet when
195 given organised tasks, instead of letting them explore the exhibits on their own. As in many science
196 centres, noise from school classes in the exhibition area can be annoying. Therefore, the learning trails
197 have been designed so that the single tasks are performed at different locations in the museum.

198 4.1. Concept of the Learning Trails

199 The learning trails are designed for self-experience in small groups of up to four students, lead by
200 the teacher. The total activity in a learning trail is meant to last less than 30 minutes. Upon arrival,
201 the students are divided into groups of up to four. Each group receives a set of experiments that the
202 members of one group shall solve together. These experiments take place in the exhibition of the
203 science centre and are related to exhibits (e.g., objects, boards, and pictures), installations, and areas
204 where the necessary ingredients for the experiment are available.

205 Each group needs at least one smart device (smartphone or tablet PC) with the possibility to use
206 Bluetooth for interaction with the beacon technology that provides the location service. The smart
207 devices are used to present tasks and extra content. Alternatively, tasks and content can be handed out
208 on paper.

209 The learning trails are compatible with the Bring Your Own Device (BYOD) paradigm, so that
210 science centres do not need to distribute these devices to the visitors. However, NTM can provide such
211 devices for school classes to avoid compatibility problems, as students might not have an own device
212 or bring devices that are incompatible with the science centre's content and services.

Table 2. List of experiments of the learning trails and reference to illustration image

learning trail	#	name	description, activity	Figure
"Forces"	#1	Pirouette	use installation, change speed through body movements	3a
	#2	Cup	cup, lace, pencil; does the cup break?	3b
	#3	C. regulator	observe model, watch video	3c
"Sound"	#1	Thunder	video: thunderstorm; count seconds: sound through tube	3d
	#2	Spoon	listen via medium air/laces: two teaspoons hit each other	3e
	#3	Vacuum bell	listen while pump makes vacuum around ringing bell	3f
"Light"	#1	Light-table	try out convex and concave lenses	3g
	#2	Letterboard	experiments with long- and short-sightedness	3h
	#3	Up-Down	observe projected image through lenses, film	3i

213 Each group performs the given tasks and experiments at the stations of the learning trail.
 214 Afterwards, all students participate in a quiz implemented with the Kahoot [26].

215 4.2. Implementation of the Learning Trails

216 So far, NTM has developed three learning trails that offer content on the physical phenomena
 217 of forces, sound, and light. Each learning trail has been implemented in two modes: I) the learning
 218 trail is paper-based; i.e., instructions to the participants are printed on laminated cards; II) the learning
 219 trail and its content are available on a device such as a smart-phone or tablet PC. This device will
 220 automatically push instructions and extra content (such as illustrating videos and audio) to the
 221 participants. The content is triggered as soon as the students approach the location of the respective
 222 experiment of the learning trail.

223 The indoor-localisation technology was implemented using beacons that are based on Bluetooth
 224 Low Energy (BLE). At each location of a learning trail experiment one beacon is placed. Actions are
 225 triggered by the proximity of the device (smartphone or tablet PC) to the respective beacon. The
 226 proximity level is classified into five zones A-E using the following thresholds: A: < 1m; B: < 2.5m; C:
 227 < 5m; D: < 7.5m; and E: above. Depending on characteristics of the installation, we assume that a
 228 participant is close by when being in Zone C, but for some exhibits Zone B is more appropriate. This
 229 can be configured per exhibit.

230 While the student groups perform the learning trails, the students' devices check the beacon
 231 proximity about once every second. Notice that too high sampling rates could drain the device for
 232 battery power.

233 4.3. The Experiments of the Learning Trails

234 Each learning trail consists of three experiments, which are performed according to a pre-defined
 235 schedule. In total, nine experiments have been developed, each of them described as a sequence of
 236 presentations, questions, and tasks to be performed by the student group. The experiments include
 237 discussions in the group to reflect on the subject of the respective experiment. Table 2 gives a short
 238 overview of these experiments, and illustrative photos taken during class visits are shown in Figure 3.

239 For the learning trail forces, the *pirouette* is an installation that can be used to explore rotation
 240 movements. The students are asked to use the installation and change rotation speed through their
 241 body movements (Figure 3a). *Cup* is an experiment where a cup is attached to a lace. Given the lace is
 242 led over a pencil and the cup is released, does it hit the ground? The students are asked to perform this
 243 experiment (Figure 3b). In a third experiment the students watch and discuss a model of a centrifugal
 244 regulator (Figure 3c).

245 For the learning trail sound, *thunder* is an experiment where the students watch the video of a
 246 thunderstorm, and they count the seconds from when they see the lightning until they notice the
 247 sound of the thunder through a long pipe (Figure 3d). *Spoon* is an experiment where the students listen
 248 to two spoons hitting each other through air and through a lace as a transmission medium (Figure 3e).



Figure 3. Examples for the experiments in the three learning trails. Forces: (a)-(c). Sound: (d)-(f). Light: (g)-(i).

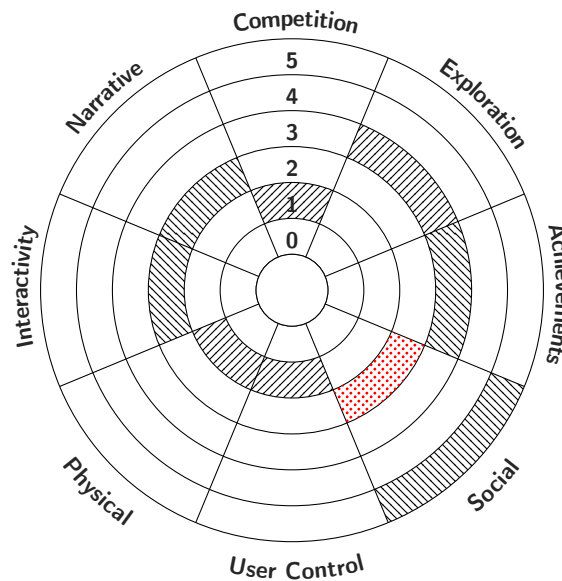


Figure 4. Engagement profile of a generic learning trail. The red dotted field indicates the S-value when the learning trail is performed as a single visitor instead of in a group.

249 In the installation *vacuum bell* a door bell is installed under a cheese dome, where a pump can generate
 250 a vacuum. The students shall observe when they stop hearing the sound from the bell (Figure 3f).

251 For the learning trail light, *light table* lets the students try out convex and concave lenses (Figure 3g),
 252 while *letterboard* provides experiments with long- and shortsightedness (Figure 3h). In *up-down*, the
 253 students observe a projected image through a set of lenses (Figure 3i).

254 5. Engagement Profile and Media Forms of the Learning Trails

255 For the analysis of the learning trails, we consider the Engagement Profile of the generic learning
 256 trail concept separately from the Engagement Profile of each experiment. In practice, we overlay the
 257 Engagement Profile of each experiment with the Engagement Profile of the generic learning trail. We
 258 also use the above described indicator to determine the media forms of the generic learning trail.

259 5.1. Engagement Profile of the Learning Trails

260 In Figure 4, we show the Engagement Profile of a generic learning trail. The values for each of
 261 the eight Engagement Profile dimensions are determined by considering which of the phrases in the
 262 description fits best. As this process is based on subjective considerations, we note the following: Since
 263 learning trails are usually performed in groups, we set the social dimension $S=5$ (for a single visitor
 264 $S=2$). Regarding the narrative (N) dimension, the learning trails are structured sequentially without
 265 the possibility for the student to alter this. Thus, we set narrative and user control to $N=2$ and $U=1$.
 266 The remaining values are set according to which phrase of the Engagement Profile definitions (see
 267 Figure 1) fits best for each dimension.

268 5.2. Engagement Profile for the Single Experiments

269 We created the Engagement Profile for all nine of the experiments by subjectively determining
 270 the phrase from Figure 1 that fits best. The charts for the experiments on forces, sound, and light,
 271 respectively are presented in Figure 5.

272 In the context of the learning trails, the Engagement Profile of the generic learning trail can be seen
 273 as an overlay for the Engagement Profile of each experiment. Note that some of the characteristics of the
 274 generic learning trail might be dominant over the individual rating. Thus, changing the characteristics

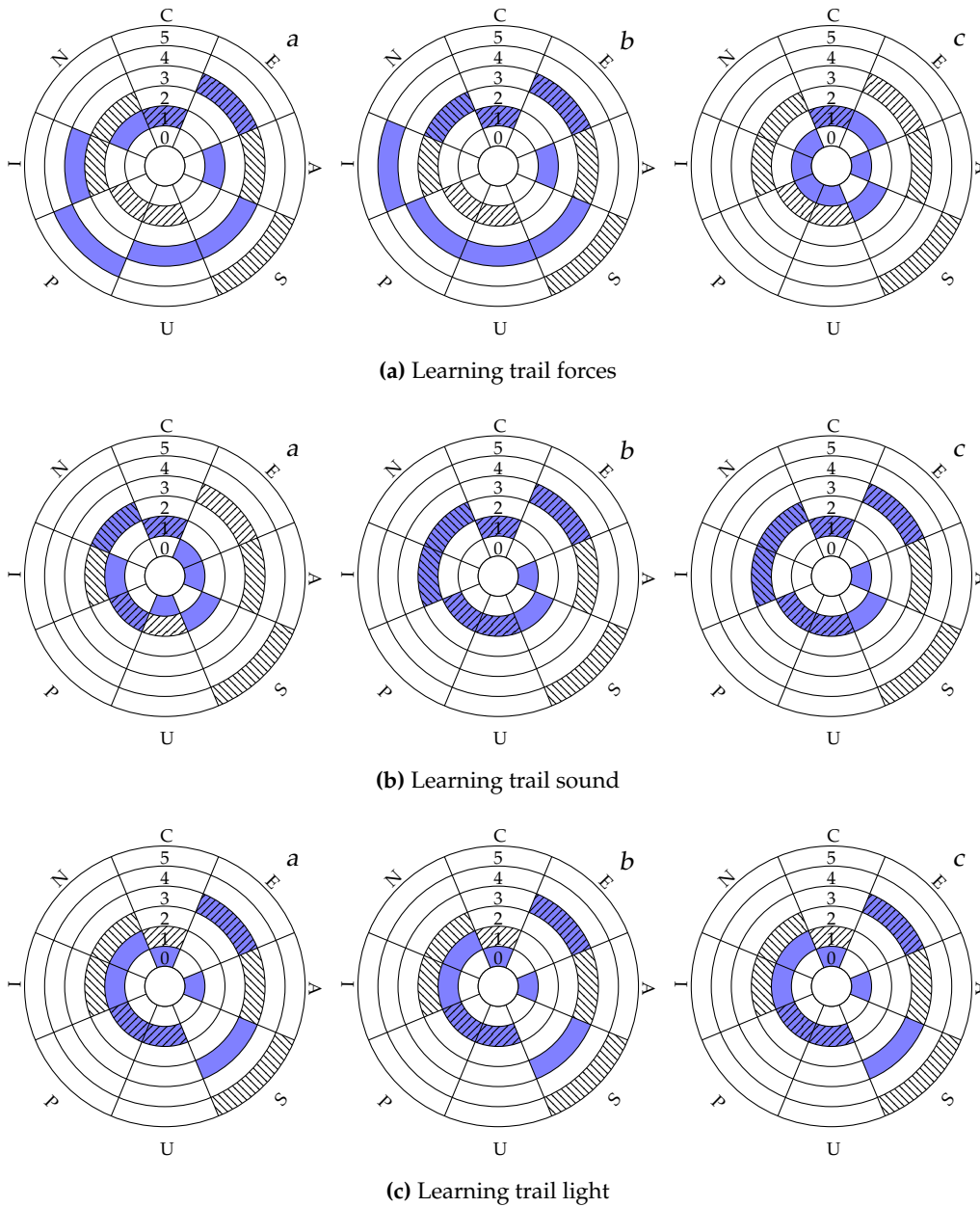


Figure 5. Engagement profile of the three experiments in the three learning trails (blue marking are for the respective experiment, hatches are for the generic learning trail).

Table 3. Media form evaluation of generic learning trails. Relevant parameters are marked with ✓ if condition is met and with ✗ else. Unmarked entries do not have an impact according to Table 1.

media form	C	N	I	P	U	S	A	E ^a	decision
Narrative	✓		✓		✓	✓ ^b		✗	✓
Interactive			✗		✗			✓	✗
Communicative						✓ ^c			✓ ^c
Adaptive	✗		✗		✗		✓	✓	✗
Productive	✗	✗	✗		✗	✓ ^c			✗

^a impact is not considered to be relevant for learning trails.

^b if visitor is alone.

^c if visitor is in group.

275 of a single experiment might not have an impact. For instance, increasing the social dimension for a
276 task might not have an impact, as the social dimension of the generic learning trail is already high.

277 5.3. Media Forms of the Learning Trails

278 Using the Engagement Profile of the generic learning trail and the translation in Table 1, we could
279 determine the applicable media forms of the learning trail concept. Table 3 shows the outcome of this
280 analysis, indicating with ✓ the conditions that applied and with ✗ those that failed. Unmarked entries
281 do not contribute to the respective media form (cf. Figure 2).

282 Note that not all values of the Engagement Profile are equally important, and considerations on
283 the impact on each value need to be made. For instance, the impact of E is considered to be weak in
284 the case of the interactive media form, as time constraints apply for school classes (i.e., the duration of
285 the visit is limited).

286 From Table 3 we conclude that the narrative and the communicative media forms apply for the
287 learning trails in their generic formation. However, when a learning trail is performed by a single
288 student, the communicative media form does obviously not apply. As a consequence, the concept
289 of the learning trails supports predominantly the activities of attending, apprehending, discussing,
290 and debating. Note that the concept of the learning trails does not focus on debating as an activity.
291 Elements of investigating, exploring, and experimenting are present, but not predominantly. The
292 activities of practising, articulating, and expressing are least present, and we recognise that the learning
293 trails are not developed for these activities.

294 6. Studying the Learning Trails

295 We wanted to explore whether we can observe differences for the operation modes *I* and *II*, as
296 well as other characteristics of the learning trails. We studied this by collecting data from school classes
297 performing the learning trail and analysed these data by aligning them with observations.

298 6.1. Test setup.

299 Each of the three learning trails consisted of three experiments, here denoted as A_t , B_t , and C_t
300 for learning trail t . After each performed experiment, the participants answered a micro survey M
301 with four questions; after the last micro-survey there was one further question denoted as survey S .
302 See Table 4 for the survey questions. Finally, all participants answered a Kahoot quiz K where the
303 correctness of the students' answers were evaluated. Thus, each group undergoes one of the sequences
304 $A_tMB_tMC_tMSK$, $B_tMC_tMA_tMSK$, or $C_tMA_tMB_tMSK$. The answers given in the micro-surveys and
305 the positioning data were stored in the respective tablet PCs and analysed later.

306 We implemented this entire procedure for both modes, that is Mode *I* for the paper-based version
307 and Mode *II* where interactive content is pushed to the students' devices when approaching the
308 respective experiment. In our study, the participants were divided into groups of three or four; one
309 participant was pointed out as the leader of the group. All group leaders received a tablet PC that was

Table 4. Formulation of the questions and scales for experiment i and learning trail t . The second column indicates the category (F=fun, R=recommend, A=use again, K=knowledge).

M_{1it}	F	How much did you like Experiment (i, t)? — scale: 1-7 (not at all – very much)
M_{2it}	R	I recommend Experiment (i, t) to others who visit the science centre. — scale: 1-7 (totally disagree – totally agree)
M_{3it}	A	When I'll visit the science centre next time, I'll use Experiment (i, t). — scale: 1-7 (totally disagree – totally agree)
M_{4it}	K	How much did you know from your school classes about the subject of Experiment (i, t). — scale: 1-5 (1=nothing, 2=a little, but don't remember much, 3=something, 4=quite a lot, 5=very much)
L		Which of the experiments A_t, B_t, C_t did you like best?

310 used for answering the micro-surveys, for logging the relative position of the device, and for accessing
 311 the content (mode *II* only). Participants that were not group leaders could answer survey questions on
 312 tablet PCs that were placed near the installations they visited. The participants were not aware of the
 313 test setup of other groups.

314 As each session included up to nine groups of students with three to four participants each, we
 315 made some precautions that groups do not interfere with each other, e.g., use the same installation
 316 concurrently.

317 We developed separate apps for each of both modes: App *I* for the paper-based version
 318 implemented the micro-surveys and logged the relative positions to the beacons. The micro-surveys
 319 were automatically triggered when a group left an area or time was up. This version did not give access
 320 to extra content. App *II* automatically presented tailored content when the participants approached an
 321 experiment; it also implemented the micro-surveys and logs. Note that all participating groups were
 322 given a tablet PC under the trials to avoid an extra bias when some groups used a tablet PC while
 323 others would not.

324 The questions of the micro-surveys given in Table 4 reflected whether the participants liked an
 325 experiment, recommendation to others, the willingness to use the experiment again, and which of the
 326 three experiments they liked best. These factors can give a good indication whether a participant liked
 327 an experiment [27]. Additionally, a self-report about the pre-visit knowledge can indicate how much
 328 they learnt, together with the number of right and wrong answers in the quiz.

329 We had planned a further question whether the participants had enough time to use the
 330 installation. However, after preliminary tests, we recognised that the learning trails were absolved
 331 much faster than anticipated. Consequently, this question was obsolete and removed from the survey.

332 6.2. Test Results

333 Students from school classes at the lower secondary school in the Oslo area participated in the
 334 study. In total, five sessions were done between Autumn 2016 and Spring 2017. In total, 113(34, 38, 41)
 335 participants appear in our log files; the numbers in parenthesis denoting participants in the learning
 336 trails “Forces”, “Sound”, and “Light”, respectively. The students were divided into groups of two
 337 to four to enable discussion and interaction in between them. One of each group was selected as
 338 spokesperson, here denoted as the group leader. The number of group leaders was $n = 41(14, 14, 13)$;
 339 $n = 14(5, 5, 4)$ for Mode *I* and $n = 26(9, 9, 8)$ for Mode *II*. The number of samples was too low to
 340 determine whether Mode *I* or Mode *II* was more engaging; we did not recognise obvious trends.

341 6.2.1. Results for all participants

342 Figure 6 shows the results from the questionnaires for the variables F , R , and A for each of the
 343 nine experiments and for all participants. As expected, the installations in the learning trails received
 344 different ratings. This result can be used by the science centre to evaluate which of the experiments are
 345 liked better than other experiments. As an observation, it seems that the leaning trail for light received
 346 a lower rating than the two other ones. Further, the experiment *centrifugal governor* received lower

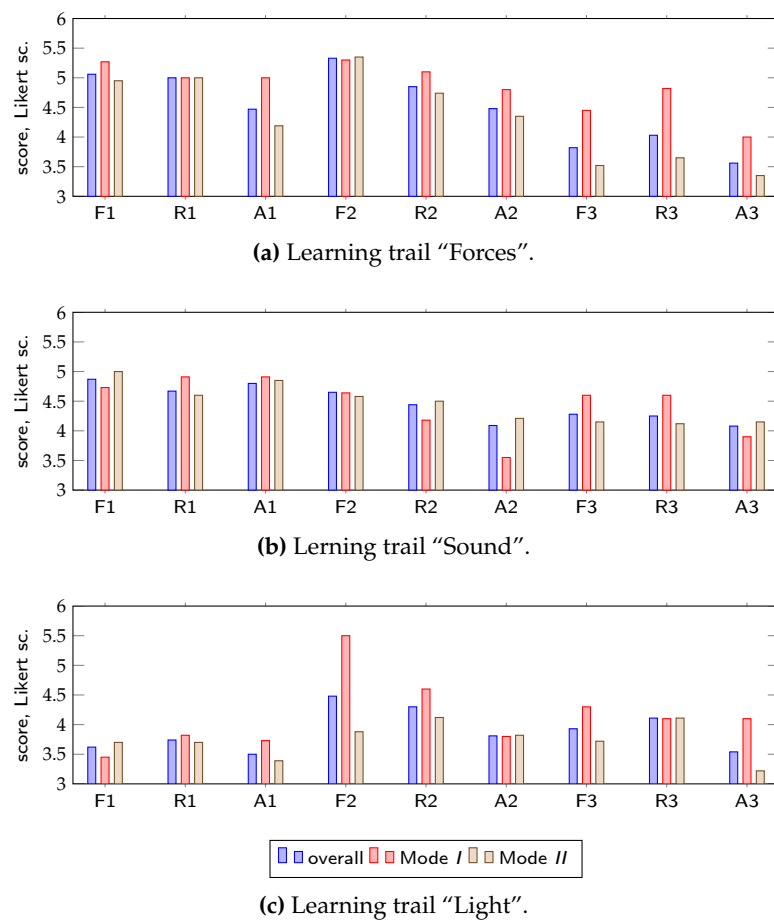


Figure 6. Response scores for tests with three learning trails and the three stations i on a Likert scale for the variables for fun F_i , recommend R_i and play again A_i . The number of participants is $n = 113(34, 38, 41)$.

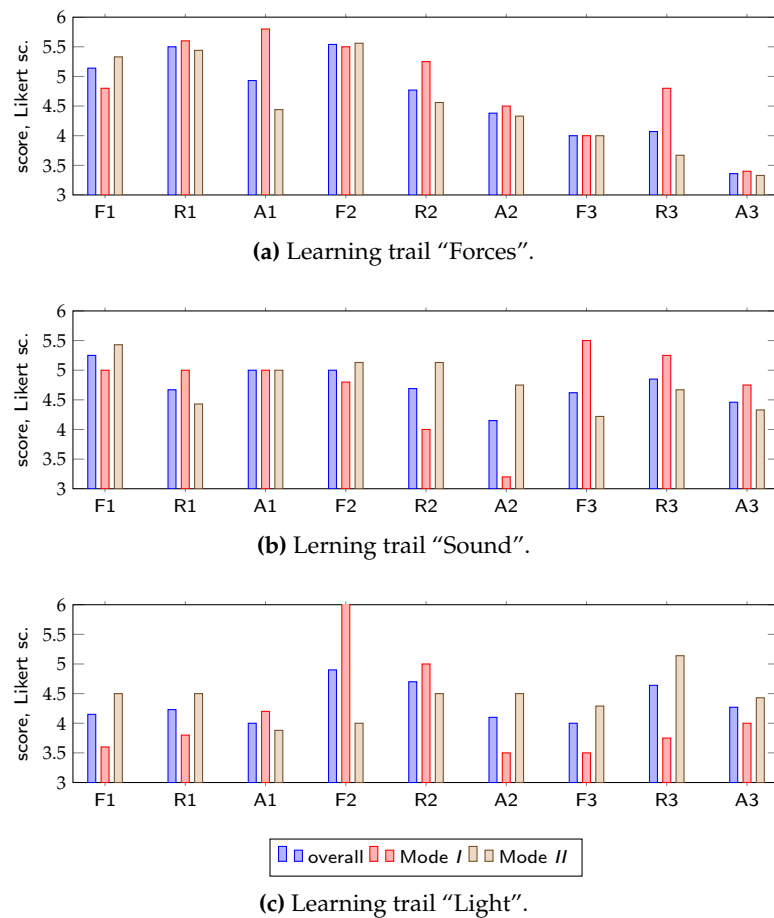


Figure 7. Group leader response scores for tests with three learning trails and the three stations i on a Likert scale for the variables for *fun* F_i , *recommend* R_i and *play again* A_i . The number of group leaders is $n = 41(14, 14, 13)$; $n = 14(5, 5, 4)$ for Mode *I* and $n = 26(9, 9, 8)$ for Mode *II*.

Table 5. Number of group leaders who liked which experiment best.

Experiment	#1	#2	#3	no vote
"Forces"	9	4	0	1
"Sound"	3	5	5	1
"Light"	4	0	6	3

347 rating than most of the others. We observed some unexpectedly high values for the learning trail for
 348 light in one of the experiments for Mode *I*.

349 6.2.2. Results for group leaders

350 As we suspected irregularities in the data set caused by the technical setup of stationary tablet
 351 PCs used for the micro surveys, we also extracted the data for the group leaders. The results for group
 352 leaders are shown in Figure 7. Still, it is not obvious which of the two modes was more engaging.

353 6.2.3. Group leaders vs. ordinary participants

354 We classified the answers into those from group leaders and ordinary participants for Mode *I*
 355 and Mode *II*. From these results, shown in Figure 8, we did not recognise significant differences in
 356 pre-visit knowledge between the participants. For the variables F , R , and A , we cannot see significant
 357 differences between the four groups. However, the participants in Mode *II* give slightly lower scores
 358 for all three variables.

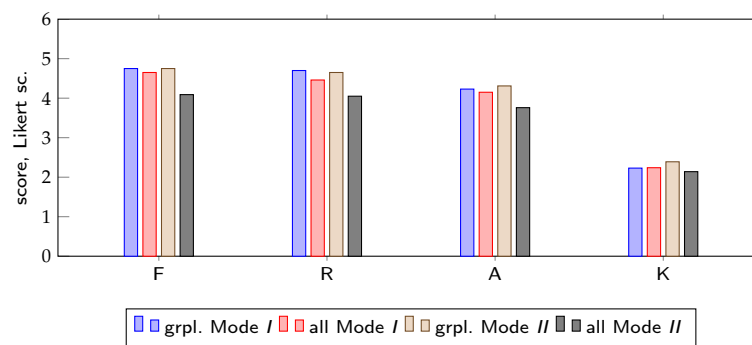


Figure 8. Response scores for tests with learning paths on a Likert scale for the variables for *fun* F, *recommend* R and *play again* A and pre-visit knowledge K.

Table 6. Results from the knowledge questions for all participants: percentage of right answers for all participants. Questions that are not related to the experiment are marked with *.

Question	"Forces"	"Sound"	"Light"
#1	73%	69%	44%
#2	21%	21%	19%
#3	24%	72%	14%
#4	9%	* 13%	11%
#5	65%	* 71%	50%
#6	* 3%	14%	88%
#7	21%	56%	0%
#8	12%	15%	42%
#9	* 56%	78%	54%
	32%	45%	36%

359 6.2.4. Experiment Preferences

360 The results for the question which of the experiments the group leaders liked best is shown in
 361 Table 5. Note that some group leaders failed to register for this question. For "Forces" and "Sound",
 362 these numbers are compatible with the results in Figure 7. However, for "Light", there is a discrepancy,
 363 as experiment #2 received no likes while it was rated rather high in the scores. As a further observation,
 364 the experiment "Pirouette" (see Figure 3a) received the highest number of mentions.

365 6.2.5. Knowledge questions

366 We evaluated the number of correct answers to knowledge questions. Each question has four
 367 alternatives where one of these is correct. Of the questions, there is always one "odd" alternative; it
 368 does not seem that the participants chose these to a large degree. Table 6 shows the percentage of
 369 correct answers. We marked questions that are about content that has not been presented during the
 370 experiment with an asterisk (*).

371 For the knowledge question, it is not significant whether Mode I or Mode II is used, nor whether
 372 the participants are group leaders. As a further observation, the pre-visit knowledge is in the average
 373 rather low (see the factor K in Figure 8).

374 6.3. Discussion of the Results

375 We went into the study with the expectation that Mode II would be preferred by the participants
 376 and, thus, resulting in higher scores. So far, we did not find evidence for this. We recognised that the
 377 number of participants in the single parts of the study is too small to show significant preferences.

378 The low impact of the mode to the result might be caused by a rather large impact of the design,
 379 activities, and use of other modalities to convey the content. In other words, the learning trail might be
 380 experienced to be multi-modal so that introducing a further mode (such as Modes I and II) has only a

381 minor impact. Another source of error might be the research setting that the students might not be
382 used to. Further, in our study, the presented content did not fully use the extra possibilities for the
383 push-medium.

384 As a note, there are other studies where the mode of presentation did not have the expected
385 impact. For instance, Vogt *et al.* [28] reported from a study where social robots were used to tutor
386 children in second language learning. They could not find the expected differences in the learning
387 effect between modes that were different in their implementation (tablet only vs. use of a social robot,
388 the latter with and without the use of gestures), although they performed a large scale study.

389 In our study, there is an indication that participants that are not group leaders in Mode II give
390 lower ratings for the variables for fun, recommend, and play again. Possibly, these participants are not
391 enough included when the group leader is working with the tablet PC. Although this effect is rather
392 small, that might be an indication that all participants should be given a tablet PC while performing a
393 learning trail.

394 There was an expectation that the students will be more quiet with organised tasks, and the
395 learning outcome will increase. During the study, we could observe that the students were more
396 quiet compared to ordinary science centre visits, although we did not perform concrete noise level
397 measurements. We leave this for future work.

398 There are indications that the characteristics of the learning trails may contribute essentially to
399 our result. The content of the learning trail for light received low scores, which can be explained by the
400 content being closer to the curriculum, being built up more theoretically, and having less engaging
401 video content than the other two learning trails. However, note that the learning outcome is not
402 necessarily related to the scores, nor to the Engagement Profile.

403 As a further note, the low scores for the experiment *centrifugal governor* could be a result of this
404 experiment consisting of looking of an object and solving a simple task. This is also visible in the chart
405 of Figure 5a.c. In contrast, the experiment *cup* (see the chart in Figure 5a.b) seems to be more engaging,
406 and will evoke more enjoyment, inspiration, creativity, activity, behaviour, and progression.

407 For NTM, the correctness level for the knowledge questions is in the usual range, compared to
408 internal studies. Commonly, the pre-visit knowledge is rather low when the students arrive at the
409 science centre. As the subjects treated by the learning trails are rather theoretical, we expected that
410 only few students were able to answer correctly. For school classes, pre-visit knowledge can often be
411 more relevant than the learning outcome from the experiments.

412 The Kahoot-quiz was performed right after the learning trails had been performed, and the learnt
413 had not yet been internalised by the students. Thus, the Kahoot could act as an engaging repetition
414 that would have helped in the internalising process of the learnt knowledge. A repetition of the Kahoot
415 some weeks after the science centre visit could have given more evidence.

416 NTM had tried several location-based concepts before; these have not worked well. In contrast,
417 the learning trails using location-based services worked well.

418 6.4. Evaluation with the Engagement Profile

419 The Engagement Profile supports design choices for improvements of installations [5], here
420 applied to the experiments of learning trails. For experiments that did not score high, we can consider
421 which dimensions of the Engagement Profile could be altered. In the context of the learning trails, the
422 social dimension S is already high, and changing S for experiments could interfere with the idea of
423 the learning trails. The factors A and E are considerably high for the generic learning trail. Thus, we
424 regard mostly the other dimensions of the Engagement Profile when considering improvements.

425 In our study, #3 of "Forces", #2 of "Sound", and all of the experiments of "Light" received low
426 scores. In Figure 5a.c, the Engagement Profile of this experiment shows rather low engagement factors
427 for #3 of "Forces". In Figure 5b.b, the Engagement Profile of this experiment shows low engagement
428 factors but E for #2 of "Sound". The same is valid for the experiments for "Light", as can be seen in

429 Figure 5c. Thus, either of the factors but S (and to some degree A and E) can be considered for an
430 improvement of these experiments.

431 From our previous work [3] we know that the factors C, P, and U are important for the target
432 group (school children). To improve the experiments of the learning trails, we recommend adding
433 more competition, adding more physical activity, and/or designing the experiments for more user
434 control.

435 6.5. Learning Trails and the GLO

436 To answer which GLO are supported by the learning trails, we revisit the five outcomes. The
437 media forms and learning experiences derived from the Engagement Profile (see Table 1) are used
438 check which of the GLO is supported by the learning trail concept.

439 **Knowledge and understanding:** The activities of attending and apprehending are supported, while
440 activities such as exploring or experimenting are more a part of the single experiment. Generally,
441 the learning trails can be said to support this outcome.

442 **Skills:** The activities of practising, but also investigating, exploring, and experimenting are not
443 predominant in the learning trails. Although these activities might be supported by the single
444 experiment, the learning trails do not support this outcome in general.

445 **Change of attitude:** The concept of the learning trails is too generic to support this outcome. However,
446 specific experiments in the learning trails might support this outcome.

447 **Enjoyment, inspiration, and creativity:** We found that the mean value of the responses for “I liked
448 the experiment” were on the positive side for almost all single experiments. Unfortunately, we
449 could not test the generic learning trail separately, as the single experiments would have had a
450 significant impact on enjoyment. For inspiration and creativity, the single experiment will have
451 most impact.

452 **Activity, behaviour, and progression:** Their outcomes are not measured with the learning trails.
453 Further, as the productive activity is not supported by the generic learning trails, this cannot be
454 said to be supported.

455 Thus, we conclude that for the concept of the learning trails, knowledge and understanding as
456 well as enjoyment are the most predominant terms of the GLO.

457 7. Conclusion

458 The science centre NTM has developed the concept of learning trails where a number of
459 thematically connected experiments are performed by a group of students. A learning trail can
460 be implemented based on laminated sheets or using apps on a tablet PC and location-based online
461 content. In our study, we could not find evidence which of these two modes is better liked or gives a
462 better learning outcome. The learning trail “Forces” performed best in our study. Further we discussed
463 the scores for the learning trails and single experiments from the findings of our study. In our study,
464 the number of samples was too low, i.e., more experiments are required to evaluate which learning
465 trail or experiment scores better than others.

466 We used the Engagement Profile as a roadmap for suggestions to improve the experiments of the
467 learning trails in the study by considering those dimensions of the Engagement Profile that are low for
468 the experiment and the overall learning trail. It is the task of designers to suggest the concrete changes
469 that are needed to increase engagement.

470 We also developed an indicator to determine the media form of exhibits. The concept of the
471 learning trails implements mostly the narrative and communicative form, supporting the activities of
472 attending, apprehending, and discussing. We also related the Engagement Profile to the media forms
473 and the GLO by setting up a set of condition for translation.

474 For NTM, the learning trail concept is promising and will be further developed to give students an
475 engaging experience in the science centre. In the study performed with secondary school students, the

476 overall rating for most of the experiments and for the concept were on the positive side, i.e., in average
477 larger than 3.5 on the Likert scale. Further, the study also gave evidence which of the experiments
478 were preferred. The Engagement Profile of the experiments can show how to adjust the experiments to
479 increase the engagement factors (see also [5]). Our research could also show that the learning trails
480 primarily support the GLO of knowledge, understanding and enjoyment.

481 **Conflict of Interest Statement**

482 The experiments were performed at the NTM, where two of the authors (J.A.A., H.H.) are affiliated.
483 The learning trails were implemented as a potential product by Expology, where one of the authors
484 (G.J.) is affiliated. Beyond these facts, the authors declare that the research was conducted in the
485 absence of any commercial or financial relationships that could be construed as a potential conflict of
486 interest.

487 **Author Contributions**

488 All authors contributed to this paper as joint work in the context of the VISITORENGAGEMENT
489 project. W.L. and I.T. developed the concept, while W.L. prepared the scientific background. J.A.A. and
490 H.H. developed the concept of the learning trails. G.J. and I.T. implemented the learning trails at the
491 NTM and performed the experiments at the NTM together with J.A.A. and H.H. W.L. and I.T. analysed
492 and discussed the data. The paper was written by W.L. with text contributions by all co-authors.

493 **Photographs**

494 The photographs in Figure 3 were taken by some of the authors during the study at NTM. The
495 persons depicted on these images have given their written consent that photographs taken during the
496 study can be used in publications.

497 **Data Protection**

498 The VISITORENGAGEMENT project is registered with the Norwegian Centre for Research Data
499 (NSD) who acts as a data protection ombud for studies performed in the project. Personal data
500 gathered during the study have been anonymised or deleted after analysis. We state that personal data
501 of participants in the study cannot be retrieved from the here published results.

502 **Funding**

503 This research has been carried out in the context of the project VISITORENGAGEMENT funded by
504 the Research Council of Norway in the BIA programme, grant number 228737. Further, this research
505 has been supported through basic institute funding at Norsk Regnesentral, RCN grant number 194067.

506 **Project Information**

507 The objective for the VISITORENGAGEMENT project was to measure the degree of engagement
508 and user experience in science centres and museums. This was done by means of sensor and camera
509 technology and the registration of user behaviour, in combination with short surveys. Project partners
510 were Expology, Norsk Regnesentral (Norwegian Computing Centre), The Norwegian Museum of
511 Science and Technology, The Norwegian Maritime Museum, Engineerium, and the Department of
512 Education at the University of Oslo. For information about the project, we refer to the project base
513 entry at the Research Council of Norway [29] and previous publications [2,3,5,11,27,30].

514 **Acknowledgments**

515 The authors wish to thank Michel de Brisis for his contributions in the research project and Anders
516 Havskjold for the implementation work of the learning trail apps and the questionnaires used during

517 the study. We also thank Ivar Solheim at Norsk Regnesentral for discussions while preparing the paper
518 and all our colleagues in the VISITORENGAGEMENT project for their input and comments.

519

- 520 1. Hofstein, A.; Rosenfeld, S. Bridging the gap between formal and informal science learning. *Studies in*
521 *Science Education* **1996**, *28*, 87–112.
- 522 2. Leister, W.; Tjøstheim, I.; Joryd, G.; Schulz, T.; Larssen, A.; de Brisis, M. Assessing Visitor Engagement in
523 Science Centres and Museums. *Journal on Advances in Life Sciences* **2016**, *8*, 49–63.
- 524 3. Leister, W.; Tjøstheim, I.; Norseng, P.G.; Joryd, G.; Bagle, E.; Sletten, H.T. Digital Storytelling
525 and Engagement in Exhibitions about Shipping. *Norsk museumstidsskrift* **2018**, *4*, 50–73.
526 doi:10.18261/issn.2464-2525-2018-02-02.
- 527 4. Hooper-Greenhill, E. Measuring Learning Outcomes in Museums, Archives and Libraries: The
528 Learning Impact Research Project (LIRP). *International Journal of Heritage Studies* **2004**, *10*, 151–174.
529 doi:10.1080/13527250410001692877.
- 530 5. Leister, W.; Tjøstheim, I.; Joryd, G.; de Brisis, M.; Lauritzsen, S.; Reisæter, S. An Evaluation-Driven Design
531 Process for Exhibitions. *Multimodal Technologies Interact.* **2017**, *1*, 1–13. doi:10.3390/mti1040025.
- 532 6. Brown, S. A critique of generic learning outcomes. *Journal of Learning Design* **2007**, *2*, 22–30.
- 533 7. Laurillard, D. *Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning*
534 *Technologies*; RoutledgeFalmer, 2002.
- 535 8. Ayudhya, W.S.N.; Vavoula, G. Mobile Family Learning in the Science Museum. Proceedings of the 16th
536 World Conference on Mobile and Contextual Learning; ACM: New York, NY, USA, 2017; mLearn 2017, pp.
537 22:1–22:8. doi:10.1145/3136907.3136948.
- 538 9. Bitgood, S. An Attention-Value Model of Museum Visitors. Report, Jacksonville State University, 2010.
- 539 10. Fredricks, J.A.; Blumenfeld, P.C.; Paris, A.H. School Engagement: Potential of the Concept, State of the
540 Evidence. *Review of Educational Research* **2004**, *74*, 59–109. doi:10.3102/00346543074001059.
- 541 11. Cooney, M.; Leister, W. Using the Engagement Profile to Design an Engaging Robotic Teaching Assistant
542 for Students. *Robotics* **2019**, *8*, 21. doi:10.3390/robotics8010021.
- 543 12. Ocampo-Agudelo, J.; Maya, J.; Roldán, A. A Tool for the Design of Experience-Centred Exhibits in Science
544 Centres. poster at Science Centre World Summit – SCWS2017, 2017. doi:10.13140/RG.2.2.22080.43520.
- 545 13. Lindauer, M. What to ask and how to answer: a comparative analysis of methodologies and philosophies
546 of summative exhibit evaluation. *museum and society* **2005**, *3*, 137–152.
- 547 14. Baldwin, T.; Kuriakose, L.T. Cheap, Accurate RFID Tracking of Museum Visitors for Personalized Content
548 Delivery. *Museums and the Web*, 2009.
- 549 15. Parsons, J.; Ralph, P.; Gallagher, K. Using Viewing Time to Infer User Preference in Recommender Systems.
550 Proc. AAAI Workshop on Semantic Web Personalization held in conjunction with the 9th National
551 Conference on Artificial Intelligence (AAAI'04); , 2004.
- 552 16. Yalowitz, S.S.; Bronnenkant, K. Timing and Tracking: Unlocking Visitor Behavior. *Visitor Studies* **2009**,
553 *12*, 47–64. doi:10.1080/10645570902769134.
- 554 17. Yoshimura, Y.; Sobolevsky, S.; Ratti, C.; Girardin, F.; Carrascal, J.P.; Blat, J.; Sinatra, R. An analysis of visitors'
555 behavior in the Louvre Museum: a study using Bluetooth data. *Environment and Planning B: Planning and*
556 *Design* **2014**, *41*, 1113–1131.
- 557 18. Moussouri, T.; Roussos, G. Examining the Effect of Visitor Motivation on Observed Visit Strategies Using
558 Mobile Computer Technologies. *Visitor Studies* **2013**, *16*, 21–38. doi:10.1080/10645578.2013.767732.
- 559 19. Moussouri, T.; Roussos, G. Conducting Visitor Studies Using Smartphone-Based Location Sensing. *J.*
560 *Comput. Cult. Herit.* **2015**, *8*, 12:1–12:16. doi:10.1145/2677083.
- 561 20. Bohnert, F.; Zukerman, I. Non-Intrusive Personalisation of the Museum Experience. Proceedings of the
562 17th International Conference on User Modeling, Adaptation, and Personalization (UMAP-09), 2009, pp.
563 197–209.
- 564 21. Bitgood, S. Designing Effective Exhibits: Criteria for Success, Exhibit Design Approaches, and Research
565 Strategies. *Visitor Behaviour* **1994**, *IX*, 4–15.
- 566 22. Mautz, R. Indoor Positioning Technologies. habil thesis, ETH Zürich, Environmental and Geomatic
567 Engineering, Institute of Geodesy and Photogrammetry, 2012.

- 568 23. Lymberopoulos, D.; Liu, J.; Yang, X.; Choudhury, R.R.; Handziski, V.; Sen, S. A Realistic Evaluation and
 569 Comparison of Indoor Location Technologies: Experiences and Lessons Learned. Proceedings of the 14th
 570 International Conference on Information Processing in Sensor Networks; ACM: New York, NY, USA, 2015;
 571 IPSN '15, pp. 178–189. doi:10.1145/2737095.2737726.
- 572 24. Bickersteth, J.; Ainsley, C. Mobile Phones and Visitor Tracking. *Museums and the Web*; , 2011.
- 573 25. Noldus, L.P.J.J.; Loke, B.; Kelia, M.; Spink, A.J. Automated Mobile User Experience Measurement:
 574 Combining Movement Tracking with App Usage Logging. *Creating the Difference: Proceedings of the Chi
 575 Sparks 2014*; van Leeuwen, J.P.; Stappers, P.J.; Lamers, M.H.; Thissen, M.J.M.R., Eds., 2014, pp. 31–34.
- 576 26. Wang, A.I. The wear out effect of a game-based student response system. *Computers & Education* **2015**,
 577 *82*, 217 – 227. doi:10.1016/j.compedu.2014.11.004.
- 578 27. Tjøstheim, I.; Leister, W.; Larssen, A.; Schulz, T. The Role of Emotion and Enjoyment for QoE – a Case Study
 579 of a Science Centre Installation. *Proc. QoMeX 2015, The 7th Intl. Workshop on Quality of Multimedia
 580 Experience*; Skodras, A., Ed. IEEE, 2015, pp. 1–6.
- 581 28. Vogt, P.; van den Berghe, R.; de Haas, M.; Hoffmann, L.; Kanero, J.; Mamus, E.; Montanier, J.M.; Oranc, C.;
 582 Oudgenoeg-Paz, O.; Garcia, D.H.; Papadopoulos, F.; Schodde, T.; Verhagen, J.; Wallbridge, C.D.; Willemsen,
 583 B.; de Wit, J.; Belpaeme, T.; Göksun, T.; Kopp, S.; Kraemer, E.; Küntay, A.C.; Leseman, P.; Pandey, A.K.
 584 Second Language Tutoring using Social Robots. A Large-Scale Study. *Proceedings of the 2019 ACM/IEEE
 585 International Conference on Human-Robot Interaction (HRI 2019)*, 2019.
- 586 29. Visitor Engagement. Entry in PROSJEKTBANKEN, Research Council of Norway, [https://www.
 587 forskningsradet.no/prosjektbanken/#/project/NFR/228737/Sprak=en](https://www.forskningsradet.no/prosjektbanken/#/project/NFR/228737/Sprak=en), 2017. accessed: 8 September,
 588 2018.
- 589 30. Leister, W.; Tjøstheim, I.; Joryd, G.; Schulz, T. Towards Assessing Visitor Engagement in Science Centres
 590 and Museums. *Proc. PESARO 2015, The Fifth Intl. Conf. on Performance, Safety, and Robustness in
 591 Complex Systems and Applications. IARIA, 2015*, pp. 21–27.

592 Abbreviations

593 The following abbreviations and names are used in this manuscript:

594 BLE	Bluetooth Low Energy
BYOD	Bring Your Own Device
EP	Engagement Profile
FM	Frequency Modulated
GLO	Generic Learning Outcomes
595 GPS	Global Positioning System
NSD	Norwegian Centre for Research Data (norw. <u>N</u> orsk <u>S</u> enter for <u>F</u> orskning <u>s</u> data)
NTM	Norwegian Museum for Science and Technology (norw. Norsk Teknisk Museum)
PC	Personal Computer
RFID	Radio Frequency IDentification
TMSI	Temporary Mobile Subscriber Identity