



HAL
open science

Principles and Validations of an Artificial Intelligence-Based Recommender System Suggesting Acceptable Food Changes

Jules Vandeputte, Pierrick Herold, Mykyt Kuslii, Paolo Viappiani, Laurent Muller, Christine Martin, Olga Davidenko, Fabien Delaere, Cristina Manfredotti, Antoine Cornuéjols, et al.

► To cite this version:

Jules Vandeputte, Pierrick Herold, Mykyt Kuslii, Paolo Viappiani, Laurent Muller, et al.. Principles and Validations of an Artificial Intelligence-Based Recommender System Suggesting Acceptable Food Changes. *Journal of Nutrition*, 2023, 153 (2), pp.598-604. 10.1016/j.tjnut.2022.12.022 . hal-03975185

HAL Id: hal-03975185

<https://hal.science/hal-03975185v1>

Submitted on 17 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Principles and validations of an Artificial Intelligence-based recommender system suggesting**
2 **acceptable food changes**

3 Jules Vandeputte¹, Pierrick Herold², Mykyt Kuslii², Paolo Viappiani³, Laurent Muller⁴, Christine
4 Martin¹, Olga Davidenko², Fabien Delaere⁵, Cristina Manfredotti¹, Antoine Cornuéjols¹, Nicolas
5 Darcel^{2*}

6

7 ¹ Université Paris-Saclay, INRAE, AgroParisTech, UMR MIA Paris-Saclay, Palaiseau, 91120, FRANCE

8 ² Université Paris-Saclay, AgroParisTech, INRAE, UMR PNCA, Palaiseau, 91120, FRANCE

9 ³ Université Paris Dauphine, Université PSL, CNRS, LAMSADE, Paris, FRANCE

10 ⁴ Université Grenoble Alpes Grenoble INP Institut d'Ingenierie et de Management, Univ. Grenoble
11 Alpes, INRAE, CNRS, Grenoble INP, GAEL, Grenoble, 38000, FRANCE

12 ⁵ Danone Nutricia Research, Centre Daniel Carasso, RD 128, Palaiseau, 91767, FRANCE

13 * corresponding author: nicolas.darcel@agroparistech.fr

14 **Short Title:** AI-based recommendations of food swaps

15

16 **Abstract**

17 **Background:** Along with the popularity of smartphones, Artificial Intelligence-based personalised
18 suggestions can be seen as promising ways to change eating habits toward more desirable diets.

19 **Objectives:** Two issues raised by such technologies were addressed in this study.. The first hypothesis
20 tested is that a recommender system based on automatically learning simple association rules between
21 dishes of the same meal would make it possible to identify plausible substitutions for the consumer. The
22 second hypothesis tested is that for an identical set of dietary swaps suggestions, the more the user is -or
23 thinks to be- involved in the process of identifying the suggestion, the higher is their probability of
24 accepting the suggestion. **Methods:** Three studies are presented in this article, firstly, we present the
25 principles of an algorithm to mine plausible substitutions from a large food consumption database.
26 Secondly we evaluate the plausibility of these automatically mined suggestions via online tests on a
27 group of 255 adult participants. Afterwards, we investigated the persuasiveness of three suggestion
28 methods of such recommendations in a population of 27 healthy adult volunteers, via a custom designed
29 smartphone application. **Results:** The results firstly indicated that a method based on automatic learning
30 of substitution rules between foods performed relatively well identifying plausible swaps suggestions.
31 Regarding the form that should be used to suggest, we found that when users are involved in selecting the
32 most appropriate recommendation for them, the resulting suggestions were more accepted (OR = 3.168
33 $p < 0.0004$). **Conclusion:** This work indicates that food recommendation algorithms can gain efficiency
34 by taking into account the consumption context and user engagement in the recommendation process.
35 Further work needs also to be done on identifying nutritionally relevant suggestions.

36 **Keywords:**

37 behaviour change; food recommendation algorithms; decision sciences

38

39 **Introduction**

40 Changing eating behaviors is critical to ensure food systems that are both healthy and environmentally
41 sustainable (1). These changes are very difficult since eating habits are firmly rooted (2) and resistance
42 to changes is often strong (3,4). The use of computer-based recommender systems appears to be a
43 promising strategy to change consumption behaviour towards more desirable diets. Smartphones have
44 become a personal assistant for many individuals, beyond simply allowing customers to search for
45 product information, compare prices, and seek feedback, smartphone apps could use algorithms to
46 provide users with personalised suggestions (5). Such algorithms would interact with humans (6) to
47 propose a series of punctual substitutions (7) or new recipes or entire meals (8) and thus move
48 individuals towards healthier diets through small iterative changes. Research in this field of application
49 is still largely uncomplete and among the many research questions to be addressed in food
50 recommendation technologies. The consumer acceptability of recommender systems relies on (i) the
51 user's compliance with the machine and (ii) the relevance of its suggestions. Therefore two questions
52 seem particularly salient: first, **what** should the artificial intelligence suggest? Second **how** should the
53 artificial intelligence suggest? This question concerns the form of human-computer interaction that is
54 most likely to ensure user compliance.

55 A first difficulty is that eating decisions are complex processes (9) influenced by numerous and
56 intricate factors, such as individual preferences, social influences or contextual effects (10–14) and they
57 all display a strong variability across individuals (15) or groups (16). Identifying acceptable
58 recommendations will therefore require an understanding of how one or several of these factors shape
59 acceptability (17,18). While the literature on how acceptability is driven (19), either by individual
60 preferences, food variability and dynamics, or by the effects of the social context is rather extensive, few
61 studies have focused on the rules governing the meal composition. A meal or a menu is not just a
62 random arrangement of foods, it is a complex assembly that often complies with very strict rules of
63 associations or exclusions between food items. Since the deduction of such rules by a human observer
64 would be very limited, machine learning algorithms could be used to explore large volumes of

65 consumption data to come up with relevant suggestions. The first hypothesis tested is that a
66 recommender system based on automatically learning simple association rules between dishes of the
67 same meal would make it possible to identify plausible and acceptable substitutions from a consumer
68 standpoint.

69 Finally, considering that the acceptability of a suggestion varies depending on how it is presented to
70 users, identifying the most persuasive recommendation method is an essential issue. Several directions
71 have so far been explored, such as personalization (20), gamification (21) but also the engagement of
72 the user in the selection process (22). The question of the effect of user engagement on acceptability is
73 an open question. Indeed, two opposite hypotheses can be formulated. It can be argued that receiving a
74 recommendation without much effort can be perceived as comfortable and attractive to the user, thus
75 favoring his or her future acceptability. Conversely, it can be assumed that if the user feels that he or she
76 has control over the suggestions, he or she will be more inclined to accept them in the end. The second
77 hypothesis tested is that for an identical set of dietary swaps suggestions, the more the user is -or the
78 more the user thinks to be- involved in the process of identifying the suggestion, the higher their
79 probability of accepting the suggestion.

80 In this article we present three studies related to testing these hypotheses. The first two studies are
81 related to our first hypothesis about the validity of automatic identification of food recommendation the
82 last work focused on the hypothesis formulated above on the most effective form of the
83 recommendation. Firstly, we present the principles of a method to reveal relevant substitutions from a
84 large food consumption database. Secondly we evaluate the plausibility of these automatically mined
85 suggestions via online tests. Afterwards, we investigate the persuasiveness of three suggestion methods
86 of such recommendations in a population of healthy volunteers.

87

88 **Methods**

89 Design and implementation of an algorithm able to identify relevant substitutions

90 Problem statement and objectives

91 Our first objective was to develop an algorithm able to mine substitutable foods. Given a database of
92 consumed meals, we aimed to extract substitutability relationships based on the implicit rules applied by
93 consumers when they compose their meals. The hypothesis on which we based our mining method of
94 relevant substitutions was that “two food items are substitutable if they are consumed in a similar dietary
95 context but rarely together”. A detailed description of this algorithm has been realised and was the
96 subject of a previous publication (23).

97 Defining context

98 The dietary context of a food item x is the set of food items C with which x is consumed. For instance, in
99 the meal (*coffee, bread, jam, juice*), the dietary context of (*coffee*) is: (*bread, jam, juice*)

100 Computing a substitutability score

101 Substitutability is not a binary relationship because foods can be substitutable to various extents. For
102 instance, it is plausible to replace potatoes with rice, less plausible to replace them with bread, but much
103 less plausible to replace those same potatoes with ice cream. Moreover, if two items are consumed
104 together, they are less substitutable because they might be associated. Therefore, we designed a function
105 to quantify the relationship of substitutability that incorporates the possibility of associativity. Detailed
106 computing of the substitutability score are presented in supplementary methods

107 Mining of relevant substitutions

108 The French dataset INCA2 (25) was used for the mining of relevant substitutions. This dataset is the
109 result of a survey conducted during 2006-2007 on individual food consumption. Individual 7-day food
110 diaries were reported for 2624 adults and 1455 children over several months accounting possible

111 seasonality in eating habits. It should be noted that since the start of this study, which was the subject of
112 an initial communication, another more recent survey has become available (26). Since the results of the
113 search for substitutability differed little over this period of time, we prefer to maintain consistency with
114 the initial study. A typical day was composed of three main meals: breakfast, lunch and dinner. The
115 moments in between were denoted as snacking. For the main meals, the location (home, work, school,
116 outdoor) and the companion (family, friends, co-workers, alone) were registered. The 1280 food entries
117 were organised in 110 groups of food items. We chose to consider this “subgroup” level of hierarchy to
118 capture substitution relationships inter-groups and intra-groups. Indeed, it would have been possible to
119 choose a finer categorisation (more groups) but this would have led to the identification of only
120 substitutions between very similar sub-groups. On the other hand, we could have chosen a coarser
121 categorisation (fewer groups), but this would have greatly constrained the algorithm's searches. The level
122 of grouping chosen was hence determined in order to obtain the best compromise between very relevant
123 but closely related changes and important but not very relevant changes. Only adults are considered in
124 this work. The meal database was split according to contextual (type of meal) information in order to get
125 better results. We compared the results of our methodology on three sub-datasets corresponding to
126 (breakfast, lunch and dinner) as well as (breakfast and lunch and dinner).

127 Evaluation of the plausibility of the substitutions mined by the recommendation algorithm

128 To evaluate the extent to which the algorithm was able to generate plausible substitutions suggestions, an
129 online task was designed in the manner of a “Turing test” (27) in which participants were asked to guess
130 whether the proposed food substitutions were issued by an artificial intelligence or by a human being.

131 Participants

132 Volunteers were invited to participate in an online experiment via a public mailing list run by the French
133 National Centre for Scientific Research (Information Relay in Cognitive Sciences, Paris, France,
134 www.risc.cnrs.fr). The inclusion criteria were to be over 18 years of age and to be able read and
135 understand French language properly. Participants could not participate more than once. Upon

136 completion of the experiment, participants could enter in a draw to win 15 €. A total of 255 participants
137 participated in the study, none of them reported to have guessed the objective of the study.

138

139 Online task

140 The experimental task consisted of three presentations of a series of 12 meals for which a proposal for
141 change was made. Proposals were made either by a professional dietician, or by the substitution mining
142 algorithm. In the latter case, to test the relevance of the substitutability scoring system, algorithm
143 suggestions either reflected substitutions with the highest substitutability score (expert algorithm), or
144 substitutions with a low substitutability score (clumsy algorithm). All these suggestions concerned the
145 same items of the 12 same meals. For each pair of meal + modified meal, participants had to answer
146 ('yes' or 'no') to the following question: "*some of these suggestions are made by an artificial*
147 *intelligence, others by a dietician, do you think this substitution was made by an artificial intelligence?*".
148 The supporting software was developed using the PC IBEX platform (28).

149 Data analysis

150 The dependant variable is the binary answer to the question on the emitter of the substitution
151 recommendation (human / non-human). Binary logistic regressions and resulting odds ratio were used to
152 evaluate whether the answer was influenced by the actual type of emitter.

153 Acceptability of identified recommendation on a group of volunteers

154 This work examined whether the extent of the user involvement in the suggestion process substitutions
155 by the AI-based recommender system affects the acceptability of that suggestion. For this purpose, a
156 food coaching interface based on the plausible substitution mining algorithm presented earlier was tested
157 on participants. Participants were different from the volunteers of the previously described online study.

158 Ethics approval

159 The study was conducted according to the Helsinki declaration guidelines and all procedures were
160 approved by the Ethics Committee of Université Paris-Saclay (decision CER-Paris-Saclay-2021-055).
161 Written informed consent was obtained from all participants. Access to the General Data Protection
162 Regulation (GDPR) is permanently available from the application interface. The participation in the
163 experiment was compensated by a gift voucher worth 50€.

164 Participants

165 Based on similar studies (18, 23–25), we estimated that 30 volunteers were needed for this study.
166 Considering a dropout and non-completion rate of the experiment of 50%, to obtain approximately 30
167 complete and exploitable responses, 60 candidates were recruited. The recruitment was done via an
168 online form distributed via a public mailing list run by the French National Centre for Scientific
169 Research (Information Relay in Cognitive Sciences, Paris, France, www.risc.cnrs.fr). The inclusion
170 criteria were being over 18 years old, not being on a diet, and owning a smartphone. To avoid possible
171 effects of the order of presentation, the participants received the different modalities in a random way.

172 Operation of the algorithm:

173 A smartphone application (virtual nutrition coach) was specifically designed for this study. The principle
174 was as follows: 1- the participant declares the meal they intend to eat the next day to the virtual coach. 2-
175 the virtual coach makes a substitution suggestion targeting one of the items of the meal according to
176 three modes of suggestion (detailed below). 3- The participant can accept or refuse the suggestion(s). 4-
177 If they accept, they commit to implement the recommendation and to certify it by sending a picture of
178 their meal via the mobile application. For 1 declared meal made by the user, 4 suggestions were
179 identified by the virtual nutrition coach via an online query to the substitution mining algorithm
180 (described above), the list of these possible substitutions was based on the 4 most substitutable items
181 presenting a better nutritional profile (according to their Rayner's score (32)). The three modalities
182 were:

183 - (A) all 4 options were presented simultaneously, and the user could choose either one or nothing

- 184 - (B) identified options were presented one by one and each time the user could refuse (in which
185 case they must justify their refusal) until the list of proposals is exhausted, if the last proposal is
186 not accepted then no option is chosen.
- 187 - (C) the coach asked the user for their preferences and proposed a single dish, which best
188 matched the announced criteria. The user could either accept or refuse.

189 Measured parameters

190 At the beginning of the experiment, the volunteers filled out a questionnaire indicating their age, sex and
191 BMI. For each recommendation session, the acceptance or refusal data were recorded, as well as the
192 constituents of the meals filled in by the volunteers and the elements suggested by the coach.

193 Data collection procedure

194 The study used a within-subject design to test two meal conditions. The experiment lasted 3 weeks,
195 spanning June and July 2021. During the sessions, the recruited volunteers had 6 recommendation
196 sessions every Tuesday and Thursday, during which they received a suggestion that they could accept to
197 implement the next day. These 6 sessions were divided into 2 sessions for each of the recommendation
198 methods.

199 Statistical analysis

200 In order to explain the influence of the mode by which the recommendation is given on the probability of
201 acceptance, a binary logistic regression analysis has been implemented. The statistical model used is
202 therefore described as follows:

$$p(\text{accept}) \sim \text{sex} + \text{method} + \text{sessionnumber} + \text{age} + \text{BMI}$$

203 All statistical analyses were performed using R (version 3.6.3) and R-Studio (RStudio 2021.09.1+372
204 "Ghost Orchid" Release). To represent the probabilities of acceptance, odds ratios were computed from
205 the logistic regression model.

206 **Results**

207 Identification of substitutions

208 Applied to the three meals datasets (breakfast, breakfast and lunch, lunch), the algorithm retrieved a
209 series of substitutable items for all considered items in the database. Substitutable items for each
210 element of a list of items for breakfast are represented in figure 1. If substitutions across food categories
211 turned out to be proposed by the algorithm, the most frequent substitutable items were intra-category
212 substitutions. When considering all foods listed in the consumption database, on average 6 % of all
213 retrieved substitutions were within the same food category, this proportion increased to 20 % when only
214 considering the three most substitutable items for each food. The substitutions proposed were also
215 consistent with regards to eating practices: substitutes of drinks were also drinks: 54 % if considering all
216 substitutions; 100 % when considering three most substitutable items (e.g.: the substitutes of coffee are
217 tea, cocoa and chicory) or the substitutes for butter for breakfast are spreadable items (26 % vs 100 %).
218 Detailed description of the results obtained by this algorithm and in particular of the comparisons
219 between substitutability scores according to the meals was carried out and was the subject of a previous
220 publication of our team (23).

221 *INSERT FIGURE 1 HERE*

222 Plausibility of mined substitutions retrieved by the algorithm

223 A total of 255 participants participated in the study, none of them reported to have guessed the objective
224 of the study. When comparing human and artificial intelligence recommenders, we found that the
225 probability that participants judge recommendations made by a human to be "made by non-humans" was
226 low (0.26 ± 0.01), while the probability that participants judge recommendations made by an artificial
227 intelligence (clumsy and expert) to be "made by non-humans" (0.67 ± 0.01) this difference appeared as
228 highly statistically significant ($p < 10^{-16}$). When comparing the two artificial intelligence-based
229 emitters (clumsy vs expert), we observed that a recommendation made by the "clumsy" artificial
230 intelligence had a significantly higher chance of being judged as not emitted by a human (0.71 ± 0.01)

231 than a recommendation made by the "expert" (0.64 ± 0.01), here again this difference appeared as
232 highly statistically significant ($p < 10^{-10}$).

233 Real life acceptability of identified suggestions and effect of recommendation methods

234 Of the 60 candidates initially enrolled in the experiment, the results of 27 of them were complete and
235 exploitable for analysis (yielding a total of 162 meals). The final sample was composed of 20 women
236 and 7 men. The average age was 37.5 ± 15.2 years. The average BMI was 22.2 ± 4.1 with two
237 overweight individuals and two others moderately obese. Of the 162 interaction outcomes between the
238 participants and the coach we observed that 74 of them resulted in the acceptance of a recommendation,
239 reflecting an overall average acceptability of 46%.

240 Analysis of the odds ratios corresponding to the different factors of influence of the acceptability
241 revealed that modality B only, (*i.e.* when the user is participating to the selection of the suggestions, OR:
242 3.168; CI 95%: 1.688, 6.061; $p < 0.0004$) was associated with an odds ratio exceeding the significance
243 threshold. A tendency was noted for the effect of age ($p=0.08$) indicating that younger participants had a
244 higher propensity to accept recommendations. Sex ($p=0.14$), and BMI ($p=0.32$) did not have a significant
245 effect on acceptability.

246

247 **Discussion**

248 This work presents an artificial intelligence-based food recommendation system designed to make
249 suggestions of food substitutions to its users. Two critical issues raised by such a recommendation
250 technology were addressed in this study: first, the method to identify relevant substitutions suggestions
251 and, second, the most efficient form that should be used to ensure acceptability from users. To identify
252 acceptable substitutions, we hypothesised that two foods consumed with the same other foods can be
253 replaced by each other with a relatively high probability of acceptance. Our results indicated that such a
254 method, based on the analysis of a large-scale food consumption database and without any prior
255 information of the considered dishes, showed overall coherence as the suggested swaps were often

256 within similar food types. We also found that such mining of relevant substitutions based solely on the
257 analysis of the meal context, allowed to produce plausible suggestions from a human perspective.
258 Regarding the form that should be used to suggest a substitution, our results obtained from a group of 27
259 volunteers over a period of 3 weeks suggest that, when users are involved in the selection of the
260 suggestion, the resulting suggestions are more likely to be accepted.

261 The substitutability score presented in this study was based on the estimation, for two foods to be
262 substituted, to what extent these items are consumed in similar meals (i.e., with the same other foods).
263 While no semantic information describing ingredients or usual positioning of foods in meals was
264 available for the recommender system, substitutions between food items of the same nutritional food
265 groups were found. Such contextual information appears to be relevant to derive food substitutions. To
266 our knowledge, this is the first study showing the richness of the information contained not in what we
267 eat but in what we combine with what we eat. A substantial amount of research has already been
268 performed in the area of recommendations to induce behaviour changes (7,8,18,33), most existing
269 approaches focused on recommending similar foods (with similar taste for instance) to consumers
270 without considering any additional contextual information. By showing that such information may
271 matter in the acceptability decision, we believe that we are opening a promising field of research for
272 automated dietary recommendations. The main strength of this study is that it links fundamental
273 algorithmic considerations related to dietary decisions with studies on healthy volunteers in online or
274 real-life consumption conditions.

275 Nevertheless, this work has focused on identifying solutions that are acceptable to the consumer. In this
276 context, we have chosen to study the plausibility of the identified solutions only from the consumer's
277 point of view, without trying to identify changes leading to a better nutritional quality diet. But it is
278 obvious that such technologies, in order to be really effective on improving the quality of diets, will have
279 to integrate additional filtering rules allowing to select, among the plausible suggestions, only the foods
280 improving the quality of the diet. For this purpose, scores taking into account past food consumption
281 history could be used (34).

282 Additional limitations are that food swaps recommendations issued by this work cannot be generalised
283 as the data set was collected in France and may not be relevant for other countries. Also as the dataset
284 was obtained in 2006-7, the substitution relationships are likely to have changed along with the
285 modification of the food offer, for example, an increase in plant-based meat substitutes would require
286 new data on their consumption relationships. Future work will need to address these limitations.

287 In a second step, using a task in which participants were invited to judge if the emitter of a given
288 suggestion was a human or a machine, we observed that recommendations made by an artificial
289 intelligence were often recognized as originating from a non-human recommender. This indicates that
290 there is a margin of progress in our ability to derive substitutability information from food consumption
291 databases. Interestingly, we observed that depending on the substitutability scores chosen by the artificial
292 intelligence for selecting the recommendation (best scores or weaker scores), the participants estimated
293 differently the plausibility of the resulting suggestions, as the higher the score was the lower the
294 probability of judging the emitter as a non-human. This substitutability score is thus a promising but
295 perfectible proxy of the acceptability of substitutions. To gain predictive capacity, it could be useful to
296 extend the concept of “contextual information” to include other information that is known to influence
297 eating decisions such as the consumptions made at the preceding meals, the time and/or the location
298 (9,35) of the meal, or the company of others (36,37).

299 It is however important to be cautious when interpreting these results. We indeed assumed that the
300 criteria on which participants based their response was the plausibility of the suggestion (e.g. “*if I think it*
301 *is a non-sense, it must have been issued by a machine*”). However, it cannot be excluded that
302 participants' responses were based on other criteria. The reasoning that they would attribute to the
303 machine (e.g., “*the recommender is expected to make recommendations being of higher nutritional*
304 *quality, if it is not the case it must be a machine*”) could strongly influence the results. It would thus be
305 interesting to combine these approaches with qualitative measures in which explanations of the reasons
306 for the choices indicated in the task would be recorded.

307 In a third step, we showed on a cohort of 27 participants that among three modalities of suggestion
308 presentation, the one that seemed to have the highest efficiency was the one that involved the user in
309 repeated exchanges, and even if the suggested options were the same as what would have been offered
310 by the other modalities, the impression of dialogue and thus of control over the production of the
311 recommendation yielded the greater acceptability from the consumer. In the field of recommender
312 systems, such an approach is called critiquing-based recommendation as it relies on users' feedback
313 (critiques) to iteratively improve the recommendation's acceptability. The advantage of such an
314 approach can be explained by the illusion of control that individuals may have in the case of a critiquing-
315 based system. Indeed, studies in psychology or behavioural economics on the endowment effect (38)
316 (popularly known as the IKEA effect, (39)) have established that the more we engage in a task (in this
317 case, finding a consumption suggestion), the more we are attached to the result, and the greater the value
318 of this result is for us.

319 The critiquing-based recommendation modality system has emerged as the most acceptable, it is
320 interesting to note that this performed even better than modality C which registered preferences before
321 making a recommendation. This advantage could be explained by the fact that when user's preferences-
322 based recommender systems have little information about the users they often fail to establish a
323 meaningful profile (cold start problem (40)). Conversely, it is likely that if virtual coaches had access to
324 large amounts of data on user preferences, they would be more effective than critiquing-based coaching
325 modalities.

326 Such results should however not be generalised, as the profile of the participants remains not
327 representative, being constituted mainly by young women. The number of participants and the design of
328 the study did not allow to identify different profiles, but, if the modality based on "critiquing" seems to
329 be more effective, we cannot exclude that another modality would be more effective in inducing
330 behavioural changes for some subgroups.

331 **Conclusion**

332 While computerised recommender technologies may be a promising tool for changing eating behaviours,
333 several questions remain open. First, concerning the methods for selecting acceptable alternatives, we
334 proposed a substitutability score based on a limited amount of contextual information (i.e., foods eaten
335 with the food to be substituted) and showed that it can be used to identify plausible and acceptable
336 alternatives. This is a promising approach, but it is highly perfectible and will gain in quality if we
337 consider other contextual elements in the substitution retrieval method. Second, the interactions between
338 the recommendation technologies and the users are another facet determining the acceptability that it is
339 necessary to consider. We have highlighted that approaches involving the users seem to have an
340 advantage, but this work remains to be refined by putting into perspective the individual profiles and the
341 most effective suggestion modalities. Future research should therefore focus on understanding the
342 dynamics of consumption (recent food history, for example) and on the effects of the diversity of
343 consumer profiles on the types of human-machine interactions to consider. Finally, additional work will
344 also have to take into account the nutritional quality of the suggested swaps to improve overall diet
345 quality.

346

347 **Acknowledgements:**

348 The authors of this article are thankful to the "PennController for IBEX" team for providing us with this
349 user-friendly virtual experimentation domain: Florian Schwarz and specially Jeremy Zehr who answered
350 all our questions regarding the programming and coding parts of the online experiment. **Author**
351 **contributions:** J.V: methodology, investigation, formal analysis, visualization, writing – original draft;
352 P.H: methodology, investigation; M.K: methodology, investigation; P.V: conceptualization,
353 methodology, validation; L.M: conceptualization, methodology, validation; Ch.M: conceptualization,
354 methodology, validation; O.D: methodology, conceptualization; Cr.M: conceptualization, methodology,
355 validation; A.C: supervision, conceptualization, methodology, formal analysis, writing – original draft.
356 N.D: supervision, project administration, funding acquisition, conceptualization, methodology, formal
357 analysis, writing – original draft. All the authors participated in writing – reviewing and editing.
358 **Declaration of interest:** the authors declare no conflict of interest. **Funding:** This study was funded by
359 the French National Agency for Research (ANR) [[ANR-18-CE21-0008](#)]. **Sources of Support**
360 **statement:** Support to this study is exclusively from governmental entities.

361 **References**

- 362 1. Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D,
363 DeClerck F, Wood A, et al. Food in the Anthropocene: the EAT–Lancet Commission on healthy
364 diets from sustainable food systems. *The Lancet* 2019;393:447–92.
- 365 2. Jastran MM, Bisogni CA, Sobal J, Blake C, Devine CM. Eating routines. Embedded, value based,
366 modifiable, and reflective. *Appetite* 2009;52:127–36.
- 367 3. Bouton ME. Why behavior change is difficult to sustain. *Prev Med* 2014;68:29–36.
- 368 4. John LK, Loewenstein G, Troxel AB, Norton L, Fassbender JE, Volpp KG. Financial incentives
369 for extended weight loss: a randomized, controlled trial. *J Gen Intern Med* 2011;26:621–6.
- 370 5. Piotrowicz W, Cuthbertson R. Introduction to the Special Issue Information Technology in Retail:
371 Toward Omnichannel Retailing. *Int J Electron Commer* 2014;18:5–16.
- 372 6. Oh YJ, Zhang J, Fang M-L, Fukuoka Y. A systematic review of artificial intelligence chatbots for
373 promoting physical activity, healthy diet, and weight loss. *Int J Behav Nutr Phys Act* 2021;18:160.
- 374 7. Ge M, Elahi M, Fernaández-Tobías I, Ricci F, Massimo D. Using Tags and Latent Factors in a
375 Food Recommender System. *Proceedings of the 5th International Conference on Digital Health*
376 2015. [Internet] New York, NY, USA: Association for Computing Machinery; 2015. p. 105–12.
377 <https://doi.org/10.1145/2750511.2750528>
- 378 8. Freyne J, Berkovsky S. Intelligent food planning: personalized recipe recommendation.
379 *Proceedings of the 15th international conference on Intelligent user interfaces*. New York, NY,
380 USA: Association for Computing Machinery; 2010. p. 321–4.
381 <https://doi.org/10.1145/1719970.1720021>
- 382 9. Köster EP. Diversity in the determinants of food choice: A psychological perspective. *Food Qual*
383 *Prefer* 2009;20:70–82.
- 384 10. Sobal J, Bisogni CA. Constructing Food Choice Decisions. *Ann Behav Med* 2009;38:37–46.
- 385 11. Rozin P. The socio-cultural context of eating and food choice. In: Meiselman HL, MacFie HJH,
386 editors. *Food Choice, Acceptance and Consumption*. Boston, MA: Springer US; 1996. p. 83–104.

- 387 https://doi.org/10.1007/978-1-4613-1221-5_2
- 388 12. Higgs S, Ruddock H, Darcel N. People in context—The social perspective. In Meiselman HL
389 editor. *Context: the effects of environment on product design and evaluation*. Duxford, United
390 Kingdom: Woohed Publishing; 2019. p. 19–38.
- 391 13. Robinson E, Higgs S. Liking food less: the impact of social influence on food liking evaluations in
392 female students. *PloS One* 2012;7:e48858.
- 393 14. Cruwys T, Bevelander KE, Hermans RCJ. Social modeling of eating: A review of when and why
394 social influence affects food intake and choice. *Appetite* 2015;86:3–18.
- 395 15. Schüz B, Schüz N, Ferguson SG. It’s the power of food: individual differences in food cue
396 responsiveness and snacking in everyday life. *Int J Behav Nutr Phys Act* 2015;12:149.
- 397 16. Bennett G, Bardon LA, Gibney ER. A Comparison of Dietary Patterns and Factors Influencing
398 Food Choice among Ethnic Groups Living in One Locality: A Systematic Review. *Nutrients*
399 2022;14:941.
- 400 17. McAuley J, Pandey R, Leskovec J. Inferring Networks of Substitutable and Complementary
401 Products. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge*
402 *Discovery and Data Mining*. [Internet] New York, NY, USA: Association for Computing
403 Machinery; 2015. p. 785–94. <https://doi.org/10.1145/2783258.2783381>
- 404 18. Zheng J, Wu X, Niu J, Bolivar A. Substitutes or complements: another step forward in
405 recommendations. *Proceedings of the 10th ACM conference on Electronic commerce*. [Internet]
406 New York, NY, USA: Association for Computing Machinery; 2009. p. 139–46.
407 <https://doi.org/10.1145/1566374.1566394>
- 408 19. Symmank C, Mai R, Hoffmann S, Stok FM, Renner B, Lien N, Rohm H. Predictors of food
409 decision making: A systematic interdisciplinary mapping (SIM) review. *Appetite* 2017;110:25–35.
- 410 20. Aldenaini N, Orji R, Sampalli S. How Effective is Personalization in Persuasive Interventions for
411 Reducing Sedentary Behavior and Promoting Physical Activity: A Systematic Review. *Adjunct*
412 *proceedings of the 15th int. conference on Persuasive Technology*. Persuasive. 2020.

- 413 21. Radhakrishnan K, Toprac P, O’Hair M, Bias R, Kim MT, Bradley P, Mackert M. Interactive
414 Digital e-Health Game for Heart Failure Self-Management: A Feasibility Study. *Games Health J*
415 2016;5:366–74.
- 416 22. Reilly J, McCarthy K, McGinty L, Smyth B. Dynamic Critiquing. In: Funk P, González Calero PA,
417 editors. *Advances in Case-Based Reasoning*. [Internet] Berlin, Heidelberg: Springer Berlin
418 Heidelberg; 2004. p. 763–77. http://link.springer.com/10.1007/978-3-540-28631-8_55
- 419 23. Akkoyunlu S, Manfredotti C, Cornuéjols A, Darcel N, Delaere F. Investigating substitutability of
420 food items in consumption data. *Second International Workshop on Health Recommender Systems*
421 *co-located with ACM RecSys*. 2017. <https://hal.archives-ouvertes.fr/hal-02482142>
- 422 24. Jaccard P. The Distribution of the Flora in the Alpine Zone.1. *New Phytol* 1912;11:37–50.
- 423 25. ANSES. Données de consommations et habitudes alimentaires de l’étude INCA 2 - data.gouv.fr.
424 [https://www.data.gouv.fr/fr/datasets/donnees-de-consommations-et-habitudes-alimentaires-de-](https://www.data.gouv.fr/fr/datasets/donnees-de-consommations-et-habitudes-alimentaires-de-letude-inca-2-3/)
425 [letude-inca-2-3/](https://www.data.gouv.fr/fr/datasets/donnees-de-consommations-et-habitudes-alimentaires-de-letude-inca-2-3/)
- 426 26. Dubuisson C, Dufour A, Carrillo S, Drouillet-Pinard P, Havard S, Volatier J-L. The Third French
427 Individual and National Food Consumption (INCA3) Survey 2014-2015: method, design and
428 participation rate in the framework of a European harmonization process. *Public Health Nutr*
429 2019;22:584–600.
- 430 27. Turing AM. Computing Machinery And Intelligence. *Mind* 1950;LIX:433–60.
- 431 28. Schwarz F, Zehr J. PennController for Internet Based Experiments (IBEX). *Open Science*
432 *Framework*; 2018; <https://osf.io/md832/>
- 433 29. Ipjian ML, Johnston CS. Smartphone technology facilitates dietary change in healthy adults.
434 *Nutrition* 2017;33:343–7.
- 435 30. Lee J-E, Song S, Ahn JS, Kim Y, Lee JE. Use of a Mobile Application for Self-Monitoring Dietary
436 Intake: Feasibility Test and an Intervention Study. *Nutrients* 2017;9:E748.
- 437 31. Ahn Y, Bae J, Kim H-S. The development of a mobile u-Health program and evaluation for self-
438 diet management for diabetic patients. *Nutr Res Pract* 2016;10:342–51.

- 439 32. Rayner M, Scarborough P, Boxer A, Stockley L. Nutrient profiles: Development of Final Model
440 Final Report. 2005; <http://www.food.gov.uk/multimedia/pdfs/npreportsept05.pdf>
- 441 33. Vandeputte J, Cornuéjols A, Darcel N, Delaere F, Martin C. Coaching Agent: Making
442 Recommendations for Behavior Change. A Case Study on Improving Eating Habits. 2022;9.
443 <https://dl.acm.org/doi/10.5555/3535850.3535994>
- 444 34. Chaltiel D, Adjibade M, Deschamps V, Touvier M, Hercberg S, Julia C, Kesse-Guyot E.
445 Programme National Nutrition Santé – guidelines score 2 (PNNS-GS2): development and
446 validation of a diet quality score reflecting the 2017 French dietary guidelines. *Br J Nutr*
447 2019;122:331–42.
- 448 35. Bisogni CA, Connors M, Devine CM, Sobal J. Who we are and how we eat: a qualitative study of
449 identities in food choice. *J Nutr Educ Behav* 2002;34:128–39.
- 450 36. Higgs S. Social norms and their influence on eating behaviours. *Appetite* 2015;86:38–44.
- 451 37. Vartanian LR, Spanos S, Herman CP, Polivy J. Modeling of food intake: a meta-analytic review.
452 *Soc Influ* 2015;10:119–36.
- 453 38. Morewedge CK, Huh YE, Vosgerau J. Thought for Food: Imagined Consumption Reduces Actual
454 Consumption. *Science* 2010;330:1530–3.
- 455 39. Norton MI, Mochon D, Ariely D. The IKEA effect: When labor leads to love. *J Consum Psychol*
456 2012;22:453–60.
- 457 40. Lika B, Kolomvatsos K, Hadjiefthymiades S. Facing the cold start problem in recommender
458 systems. *Expert Syst Appl* 2014;41:2065–73.
- 459

460 **FIGURE 1: Substitutability scores computed between most frequent breakfast items:** The breakfast
461 sub-dataset was extracted from the French dataset INCA2 (25) gathering individual 7-day food diaries
462 for 2624 adults and 1455 children.