The development and validation of a toolkit to measure consumer trust in food

Document Version:
Peer reviewed version

The development and validation of a toolkit to measure consumer trust in food

Tony Benson [a], Fiona Lavelle [a], Michelle Spence [a], Christopher T. Elliott [a], Moira Dean [a]

Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, Belfast BT9 5AG, UK. t.benson@qub.ac.uk, flavelle01@qub.ac.uk, m.s.spence@qub.ac.uk, chris.elliott@qub.ac.uk, moira.dean@qub.ac.uk

Correspondence: moira.dean@qub.ac.uk, Tel: +442890 976561

Abstract: Consumer trust is an important aspect in the functioning of any market but particularly the food and drinks sector. Food safety incidents and changes in food production practices have simultaneously led to a decrease in consumer trust and a need for greater trust. Previous research has developed items to measure consumer trust in food, however, these have not always been subject to validity and reliability testing and there exists no collated toolkit or collection of items to measure trust in various aspects of the food system. Therefore, the current set of studies aimed to develop a valid and reliable consumer trust toolkit which can be used to measure trust in specific aspects of the food system. Study 1 consisted of a literature review of previous consumer trust measures to construct an initial toolkit of items, followed by an exploratory factor analysis (n = 481) to identify the structure of the toolkit. Study 2 (n = 1,027) used confirmatory factor analysis to verify the factor structure of the model from study 1 with six different factors (types of trust): Organisation trust, product trust, interpersonal trust, trust in the food chain, organisation distrust, and general distrust. Study 2 also established the validity of the toolkit (face validity, convergent validity, and discriminant validity). Study 3 resampled a collection of individuals from study 2 (n = 247) to establish composite reliability and temporal stability (test-retest reliability). The resultant consumer trust toolkit provides a valid and reliable collection of items which can be used in future research to measure consumer trust in selected aspects of the food system.

Keywords: Consumer, trust, food, measure, validity, reliability

1. Introduction

Consumer trust is a vital component in any market, without which the selling and purchasing of goods and services as well as the development of new products would prove difficult or impossible (Nuttavuthisit & Thøgersen, 2017). The concept of trust is broad and overlaps multiple disciplines including economics, psychology, and sociology. Such wide scope has led to the creation of various definitions of trust. For example, it has been defined as the "willingness to rely on an exchange partner in whom one has confidence" (Moorman, Deshpande, & Zaltman, 1993, p. 82) and "confidence in an exchange partner's reliability and integrity (Morgan & Hunt, 1994, p. 23) (for further definitions of trust see Bozic (2017)). Arguably the most well-known definition refers to trust as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer, Davis, & Schoorman, 1995, p. 712). With regards to the consumer, this means that they will spend based upon the expectation that the product or service that they will receive is authentic and genuine. Extensive research has identified several components of consumer trust. At a basic level, expertise and trustworthiness have been highlighted as key factors (Frewer & Miles, 2003). However, further components are more consistently found in the literature. Specifically, competence, benevolence, integrity, openness, and honesty are some of the components that have been suggested to form trust (For further details see Connolly & Bannister (2007)).

While consumer trust is essential to any market, it is particularly relevant to the food market. Consumers expect foods available for purchase to be safe and of satisfactory quality. If a consumer trusts and therefore unknowingly purchases and consumes an inauthentic or unsafe product, this may lead to consequences ranging from a poor sensory experience through to illness or death. More than ever, consumers want to know the source of their food and how it was grown, handled, shipped, produced, and packaged, with traceability and transparency key trends (Lu, Wu, Wang, & Xu, 2016). Some of this desire for information has been a consequence of various well-documented food scandals, for example, the Bovine Spongiform Encephalopathy outbreak in the 1990s and the 'horsemeat scandal' in 2013 where beef was substituted with horse meat in a number of products. Such incidents have eroded consumer trust in food (Coveney et al., 2015; Zachmann & Østby, 2011) and decreased sales in affected products (MacLeod, 2013; Miran & Akgüngör, 2005; Roosen, Lusk, & Fox, 2003; Schlenker & Villas-Boas, 2009). In addition to consumer trust being impacted by food safety incidents, trust can also influence the success of particular brands or products in the food sector. Bruschi and colleagues (2015) found that trust in EU food certification schemes may explain the success of foreign products in particular markets.

67 Trust at a wider level is also important in the food sector. Mutual trust between businesses is

68 necessary for food supply chains to operate and be successful (Meixner et al., 2009).

69 Business to business trust is essential given that those in businesses face the same issues

70 as consumers, such as the inability to scrutinize all quality characteristics of food (Fritz,

71 Martino, & Surci, 2008).

72 Current consumer trust in the food chain and production system is relatively low

73 (Coveney et al., 2015). In addition to the impact of food scares and safety incidents, it has

74 been suggested that the complexity of the food industry may also affect trust (Giampietri,

75 Verneau, Del Giudice, Carfora, & Finco, 2018). Increasing sophistication and globalisation of

76 the food market means that consumers are more distanced from the source than ever. This

77 is both literally in terms of urban living and physical proximity to farms and metaphorically

78 with regards to position amongst multiple actors in the food production chain (Berg, 2004;

79 Wilson et al., 2016). This increase in complexity and distance may not only have contributed

80 to a decline in trust but has also simultaneously meant that trust is more important for the

81 consumer than ever. Food safety and quality have been deemed as a 'black box' for

82 consumers who must rely upon and place their trust in the actors involved in various stages

83 of the food chain (De Jonge, Van Trijp, Jan Renes, & Frewer, 2007; Giampietri et al., 2018).

84 Transforming consumer trust in food is a current key challenge (Giampietri et al., 2018;

85 PriceWaterhouseCoopers, 2015). Given predicted global trends such as a rapidly increasing

86 population and climate change and the ensuing scarcity of resources, current consumer trust

87 must be rebuilt to ensure a future sustainable food system. Recently the most optimal

88 strategies for rebuilding consumer trust in food were identified by actors in the food system

89 as; transparency, protocols and procedure, be proactive, collaborate with stakeholders, and

90 put consumers first (Wilson et al., 2016). These strategies were also endorsed by

91 consumers (Tonkin et al., 2019). Communication following a food safety incident should

92 demonstrate care, commitment, consistency, coherence, and clarity (Hobbs, 2011).

93 Traceability and transparency are promising potential solutions to increase consumer

94 trust. For example, Japan has a mandatory beef traceability system allowing consumers to

95 trace beef through their mobile phone or a website using an identification number (Jin &

96 Zhou, 2014). Other studies have examined the use of traceability in relation to consumer

97 trust in food (Liu, Gao, Nayga Jr, Snell, & Ma, 2019; Menozzi, Halawany-Darson, Mora, &

98 Giraud, 2015; Wu, Wang, Zhu, Hu, & Wang, 2016). In order to understand if approaches

99 such as these to rebuild consumer trust have been or can be successful, it is necessary that

100 trust is assessed accurately. As previously mentioned, consumer trust is a broad concept

101 spanning numerous disciplines. Trust is a latent concept meaning that it is not directly

102 observable, as a result, numerous items and questions have been created to measure it.
103 Interpersonal trust considers trust as a personality type trait such that individuals are viewed
104 as having a disposition to trust and be trusting of others. For example, "Most people can be
105 counted on to do what they say they will do" (Rotter, 1967). Interpersonal trust has been
106 associated with Genetically Modified (GM) food choices (Ding, Veeman, & Adamowicz,
107 2012) and purchase frequency of organic food (Dumortier, Evans, Grebitus, & Martin, 2017).
108 Meanwhile, other studies have focused directly on trust in food or related areas. Examining
109 trust in GM food, Zhang and colleagues (2018) used two items based on previous research
110 and created two further items to measure trust in factors which might impact upon
111 acceptance of GM foods such as labelling and the media. Perrini and colleagues (2010)
112 measured trust in organic products alongside trust in a retailer, finding a greater level of trust
113 in organic products when sold by a socially responsible retailer. Another study adapted items
114 from the literature to measure perceived levels of care, competence, and openness in actors
115 involved in food safety (de Jonge et al., 2007). Meanwhile Lassoued, Hobbs, Micheels, and
116 Zhang (2015) measured trust in chicken in relation to brand trust and the food industry.

117   The aforementioned studies measure different types of trust using different items,
118 however, some of these items and scales lack validity and/or reliability suggesting that they
119 may not be scientifically sound. Additionally, while the broad range of items and scales to
120 measure consumer trust is encouraging, the field is fragmented and to the best of our
121 knowledge there exists no single comprehensive collection or toolkit to measure consumer
122 trust along the food chain. As such, the current set of studies aims to develop a valid and
123 reliable food consumer trust toolkit. Researchers can choose items to measure trust along
124 the food chain depending on the focus(es) of their study.

125   A series of three studies was used to develop and validate questionnaire items to
126 measure consumer trust in food. Each of the studies used an online survey to develop and
127 validate the toolkit. Study one involved a brief review of the literature and generation of the
128 initial pool of relevant items. Study two was used to confirm the findings of study one and
129 reduce and refine the number of items. Study three tested the reliability of the toolkit.

130

2. Study 1 – Critical review of the literature and scale development

133   This study commenced with a critical review of the literature to identify existing items
134 and measures relating to consumer trust in food. The following search terms were used in
135 combination across the PsycINFO, Web of Science, and PubMed databases: Consumer

trust, trust, confidence, food supply chain, food systems, food networks, measur*, tool*, scale*. A total of 9,048 articles were retrieved and assessed for relevance. Duplicate and irrelevant articles (such as those which did not contain a measurement of trust) were removed, leading to an in-depth review of 40 full text articles. A further manual search of the grey literature and publication lists of known authors in the field led to the inclusion of an additional five articles. The most commonly used items and scales were collated and critiqued by four researchers in the areas of food quality, safety, and nutrition (TB, FL, MS, MD). Items were selected based upon their frequency of use within the literature, face validity, and any other validity or reliability testing which had been conducted. These selected items were then administered to participants using an online survey (for further information, see section 2.1). This study aimed to generate an initial pool of items that could be used to measure consumer trust.

2.1 Method

Trust items

Following the critical review, those items identified as most appropriate and relevant by the researchers were included in the initial pool of consumer trust items (n = 54). All items were measured on a seven point scale ranging from 1 (strongly disagree) to 7 (strongly agree). Overall, the items spanned five levels of the food chain, ranging from the general to the specific:

- Interpersonal trust. Trust at the individual level. How trusting an individual is in general as a person. Example item: "Most people are basically good and kind".
- General organisation trust. Trust at the organisation level. How much an individual trusts a certain organisation (related to food but not involved in the food chain) in general. Example item: "<organisation> is dependable"
- Specific organisation trust. How much an individual trusts an organisation in terms of a specific area or to perform a specific task. Example item: "<organisation> are good at looking at the evidence and judging what to do".
- Food chain trust. How much an individual trusts the actors or organisations involved in food production. Example item: "<organisation/actor> has the competence to control the safety of food".
- Product trust. How much an individual trusts a specific product. Example item: "<product> is trustworthy".

169

Participants and procedure

171   Individuals were invited to participate in the survey by a research agency from their
172 online panel of UK consumers in October 2018. Individuals completed a series of screening
173 questions to assess their eligibility to take part in the study. To avoid bias, anyone working in
174 (or living in a household with anyone working in) food safety, food processing or
175 manufacturing as well as the farming, growing, wholesale or retail of food or drinks were
176 excluded. Those aged under 18 were also excluded. Quotas were applied to achieve a
177 nationally representative UK sample in terms of age, gender, and region. The final sample
178 number was 481, with individuals ranging in age from 18 to 92 years old (M = 46.64, SD =
179 17.06) (Table 1). Participants completed sociodemographic items followed by the trust items
180 for each type of trust. While we believed that that the trust items could be applied to any
181 organisation, product, or food chain actor, in this survey we used the European Food Safety
182 Authority (EFSA) as the organisation, beef burgers as the product, and food manufacturers
183 as the food chain actor. EFSA was defined to participants as an organisation which 'provides
184 independent scientific advice about food risks and food safety to the public and decision
185 makers who regulate food safety in Europe'. To ensure no missing data, a forced response
186 option was used for all items. The questionnaire took approximately 20 minutes to complete.
187 Ethical approval (10/18/BensonT) was granted by the Queen's University Belfast School of
188 Biological Sciences Research Ethics Committee. The study was conducted in accordance
189 with the Declaration of Helsinki and informed consent was obtained.

190   <Insert table 1 about here>

191

Data analysis

193   Prior to data analysis, where necessary, items were reversed coded so that a higher
194 score indicated greater trust for all items. Exploratory factor analysis (principal axis factoring)
195 with direct oblimin rotation was used. This oblique rotation was used as it was believed that
196 the factors (types of trust) would be related (Yong & Pearce, 2013). For example, if someone
197 has a high level of interpersonal trust and is a trusting person, one would expect they will
198 also have a high level of trust in other areas. Each iteration of the factor analysis was refined
199 using cut-off criteria outlined below until an optimal solution was reached. All analyses were
200 conducted using IBM SPSS Statistics v25.

2.2 Results

The results showed an excellent Kaiser-Meyer-Olkin (KMO) value of 0.95 (Kaiser, 1974) and a significant ($p < .001$) Bartlett's Test of Sphericity, indicating that the sample was adequate for factor analysis (Field, 2009). Eight factors were apparent in the data, as shown by Eigenvalues greater than 1 (Kaiser, 1960). However, one factor had no items which loaded highest upon this factor and another factor contained only two items which loaded highest upon it. In general, each factor should contain at least three items (Carpenter, 2018). Given this, the analysis was conducted again with the instruction to extract six factors only.

In order to evaluate the six factor solution and reduce the number of items, the following criteria were used: No factors with fewer than three items (Carpenter, 2018), no items which cross-loaded greater than 0.3 across factors, no items with communality less than 0.3, and no items with corrected item scale correlation less than 0.3 (Worthington & Whittaker, 2006). Typically, items with factor loadings less than 0.3 or 0.4 are also removed (Worthington & Whittaker, 2006). However, given the relatively small number of items removed using the aforementioned criteria and the need to further reduce the number of items in the scale at this stage, as well as the suggestion that the factor loading cut-off should be set as high as possible (Worthington & Whittaker, 2006), it was decided to use a more stringent minimum factor loading of 0.6. Items were removed only if they improved or did not reduce the reliability (Cronbach's Alpha) of that factor. In total, 9 items were removed, leaving 45 items.

The remaining items were subjected to a final exploratory factor analysis to ensure that the factor structure and results were acceptable following the previous modifications. All factors contained at least three items, no items cross-loaded on more than one factor, and the minimum factor loading was 0.6. In addition, the internal reliability values for each scale (Cronbach's Alpha) were above the typical cutoff value of 0.6-0.7 (Hair, Black, Babin, & Anderson, 2014), therefore, all 45 items were retained. A review of the factors and their associated items suggested that factor 1 related to trust in organisations; factor 2 related to product trust; factor 3 interpersonal trust; factor 4 trust in the food chain; factor 5 organisation distrust; and factor 6 interpersonal distrust.

2.3 Discussion

Study one collated items measuring consumer trust relating to the area of food. Following refinement, items loaded well on factors and there was little evidence of cross-

loading, suggesting defined and distinct factors. While it was hypothesised that the data would lead to five factors, six factors emerged. As expected, interpersonal trust, food chain trust, and product trust were apparent as separate factors in the data. However, specific organisation trust appeared to merge with general organisation trust, suggesting that consumers may not distinguish between the two. Unexpectedly, two distrust factors were identified in the data – interpersonal distrust and organisation distrust. The next step was to confirm these findings using a larger sample.

3. Study 2 – Confirmation and validation of factorial structure

Following the establishment of the factor structure in study 1, study 2 aimed to confirm these findings in a larger sample. A number of other scales and measures were included in this study to allow for validation testing of the trust toolkit, these are outlined in section 3.1.

3.1 Method

Trust items

The final set of 45 trust items from study 1 were included in study 2. These were unchanged, with items measured on a scale ranging from 1 (strongly disagree) to 7 (strongly agree) and spanning the six factors (types of trust) identified in study 1: trust in organisations; product trust; interpersonal trust; trust in the food chain; organisation distrust; interpersonal distrust.

Established item measuring interpersonal trust

A single item commonly used to measure interpersonal trust in the literature was included in addition to the 45 trust items: "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" (European Social Survey, 2012). This was measured on an 11 point scale (0 = you can't be too careful, 10 = most people can be trusted). This item was included to examine the convergent validity of the interpersonal trust factor.

Life satisfaction

Life satisfaction was measured with a single item (Office for National Statistics, 2015): "Overall, how satisfied are you with your life nowadays?" measured on a scale from 1

267 (not at all satisfied) to 10 (completely satisfied). This item was included to examine the

268 convergent validity of the interpersonal trust factor.

269

270 ### Food safety information seeking

271   To measure to what extent individuals were interested in seeking information relating

272 to food safety risks, participants were asked to select all types of information which they

273 would like to receive if new risks to food safety were discovered (adapted from Etienne,

274 Chirico, McEntaggart, Papoutsis, & Millstone, 2018). Examples of information sources

275 included 'general description of the risk' and 'technical or scientific details'. The number of

276 information sources that participants selected (ranging from 0 to 7) was used as a score to

277 indicate how interested the individual is in seeking information related to food safety, with a

278 higher score indicating greater interest. These items were included to examine the

279 convergent validity of the organisation trust factor.

280

281 ### Corporate distrust scale

282   General distrust in organisations was measured using items from the corporate

283 distrust scale (Adams, Highhouse, & Zickar, 2010). The most relevant items (seven items)

284 were adapted so that EFSA was defined as the example organisation/corporation, as was

285 the case with the items measuring trust in organisations. Items in this scale included 'EFSA

286 do not accept accountability for their actions' and 'EFSA intentionally deceives the public'.

287 These items were included to examine the convergent validity of the organisation trust

288 factor.

289

290 ### Frequency of buying

291   Individuals were asked approximately how often they purchase beef burgers (the

292 chosen product used for the items measuring product trust) using a scale from 1 (never) to 9

293 (once or more a day). This item was included to examine the convergent validity of the

294 product trust factor.

295

296 ### Food quality interest

297   To measure interest in food quality, three items were adapted from the general health

298 interest scale (Roininen, Lähteenmäki, & Tuorila, 1999): 'The quality of food has little impact

299 on my food choices' 'I am very particular about the quality of food I eat' and 'I eat what I like

300 and do not worry about the quality of food'. These items were selected as the most relevant

301 to measure interest in food quality, with other items in the scale measuring healthiness and

302 nutrients. Each item was measured on a seven point scale ranging from strongly disagree to

303 strongly agree. These items were included to examine the convergent validity of the product

304 trust factor.

305

306 Participants and procedure

307 As with study 1, a research agency invited potential participants from their online

308 panel of consumers to partake in the study (in November 2018). To ensure that the

309 developed trust toolkit had cross-cultural relevance, individuals from countries with varying

310 levels of trust were invited to participate. The 2018 Edelman Trust Barometer (Edelman,

311 2018) and the Eurobarometer 354 (European Food Safety Authority, 2010) were used to

312 select countries according to their levels of general trust and levels of confidence in

313 organisations related to food such as EFSA. Using these sources, the UK was chosen as a

314 country to sample as it has the lowest level of trust in general (alongside Ireland) in Europe,

315 and amongst the lowest levels of confidence in Europe for organisations related to the food

316 chain. Finland was chosen as a country high in confidence in organisations relating to the

317 food chain. Nordic countries (and particularly Finland) typically have a high level of trust for

318 issues relating to food (Jokinen, Kupsala, & Vinnari, 2012). Germany is typically amongst

319 either the lowest or highest ranking countries in Europe for confidence depending on the

320 food-related organisation and has an average level of trust in general amongst European

321 countries. Greece was chosen to be sampled as it is a Southern European country in

322 contrast to the aforementioned Northern European countries and this may be reflected in

323 differing attitudes to food. For example, individuals in Greece have high levels of concern for

324 food production and quality (European Commission, 2012)

325 Individuals under the age of 18 or those working in (or living in a household with

326 anyone working in) food safety, food processing or manufacturing as well as the farming,

327 growing, wholesale or retail of food or drinks were excluded from participation. The final

328 samples for each country were approximately representative in terms of age, gender, and

329 region (maximum +/- 8% difference between population figures and sample achieved). In

330 total, 1,027 individuals participated (UK n = 256; Germany n = 257; Finland n = 253, Greece

331 n = 253). The mean age was 46.99 (SD = 16.95, range = 18 to 85) (Table 2).

332

333

334   <Insert Table 2 about here>

335

336   Prior to the rollout of the survey, all questions and instructions were translated into the

337 primary language of each sample country by native speakers. These translations were then

338 proofread by a second native speaker and quality assured by a third linguist before being

339 confirmed by the project manager (also a trained linguist). As with survey 1, EFSA was

340 specified as the organisation for items measuring organisation trust, beef burgers were

341 specified as the product for items measuring product trust, and food manufacturers were

342 specified as the actor in the chain for items measuring trust in the food chain. Participants

343 completed sociodemographic details then the 45 trust items followed by the remaining items.

344 Each item in the survey used a forced response option to ensure no missing data. The

345 questionnaire took approximately 20 minutes to complete. Ethical approval (10/18/BensonT)

346 was granted by the Queen's University Belfast School of Biological Sciences Research

347 Ethics Committee. The study was conducted in accordance with the Declaration of Helsinki

348 and informed consent was obtained.

349

350 Data analysis

351   The initial step in data analysis involved recoding any reverse scored items. The final

352 model identified in study one was entered into IBM SPSS Amos v25 as a confirmatory factor

353 analysis model with maximum-likelihood estimation. This model was then assessed and

354 amended according to various model fit statistics. Following their review of the literature,

355 Hinkin and colleagues (1997) outline several fit statistics typically used to evaluate models.

356 - Chi-square (χ^2) – A non-significant chi-square value (p > 0.05) which is, at a
357   maximum, two or three times larger than its value divided by the degrees of freedom
358   (df) indicates that the model can be accepted
359 - Root Mean Square Error of Approximation (RMSEA) – A value of 0.05 or less is
360   optimal
361 - Comparative Fit Index (CFI) – A value of 0.90 or greater indicates that the model can
362   be accepted
363 - Normed-Fit Index (NFI) – A value of 0.90 or greater indicates that the model can be
364   accepted
365 - Tucker-Lewis Index (TLI) - A value of 0.90 or greater indicates that the model can be
366   accepted

367

368   Following the use of fit statistics and modification indices to refine and select the most

369 appropriate model, several validation tests were conducted. Face validity is used to ensure

370 that the items under each factor measure what they claim to measure 'at face value'.

371 Convergent validity shows that measures are valid by identifying a relationship with an
372 existing similar measure using correlation analysis. A further method of identifying
373 convergent validity is to examine the Average Variance Extracted (AVE) of each scale (a
374 summary indicator of convergence). An AVE value of 0.5 or more is considered acceptable
375 (Hair et al., 2014). Discriminant validity refers to a scale or measure being distinct from other
376 constructs. Discriminant validity is present when the square root of the AVE is greater than
377 the correlation between the scale and other scales in the model (Hair et al., 2014).
378 Discriminant validity is also present if the Maximum Squared Variance (MSV) is less than the
379 AVE for each scale (Rebelo-Pinto, Pinto, Rebelo-Pinto, & Paiva, 2014).

380

381 3.2 Results

382 Model refinement

383   Initial fit statistics showed that the model retained from survey 1 was acceptable (see
384 Table 3). The chi-square value was significant and greater than 2-3 times larger than the
385 χ^2/df, however, this fit statistic is known to be sensitive to large sample sizes such as that in
386 the current study. While the RMSEA value was greater than 0.05, this was still acceptable at
387 0.08 (Hooper, Coughlan, & Mullen, 2008). The CFI, NFI, and TLI were all acceptable at
388 approximately 0.90.

389   While the modification indices were examined for potential improvements to model fit,
390 no changes were made as these were not justifiable by theory or rationale. However, two of
391 the factors (organisation distrust and interpersonal distrust) consisted of only negatively
392 worded items. Past research on scale development has found that the inclusion of negatively
393 worded items in a scale or questionnaire may lead to spurious factors containing only these
394 items (X. Zhang, Noor, & Savalei, 2016). This is due to a method effect and the manner in
395 which participants respond to negative items, rather than these items representing true
396 factors. To test for methods effects, two further models (models 2 and 3) were created
397 (Table 3).

398   <Insert table 3 about here>

399

400   Model 2 was a modification of model 1 which allowed the error terms of the
401 negatively worded items to covary. This produced similar results to those of model 1. Model
402 3 was also a modification of model 1, which allowed the error terms of the positively worded
403 items to covary. The results of this model were also acceptable.

404   Overall, it can be seen that model 3 achieved better fit statistics than models 1 and 2.
405   This is perhaps unsurprising given the large number of positively worded items and therefore
406   large number of modifications made to the model to allow the error terms to covary. Fit
407   statistics provide guidance as to the acceptability of each model, however, theory and
408   parsimony are also important considerations in model selection. Given this, and that the
409   results show that the improvement in fit of model 3 over models 1 and 2 was marginal,
410   model 1 was established as the most relevant, parsimonious, and acceptable model for the
411   data. All items loaded well on their respective factors, with a minimum item loading in the
412   model of 0.6 (Figure 1). The final items can be seen in Table 4.

413

414   <Insert Figure 1 about here>

415

416   <Insert Table 4 about here>

Face validity

To ensure that the toolkit had face validity, four consumer researchers in the areas of food quality, safety, and nutrition (TB, FL, MS, MD) reviewed the final model. All items were examined to ensure that they measured what they claim to measure. It was agreed that all items were appropriate for their factor (type of trust) and that no items were too similar or measured more than one type of trust. Therefore, face validity was established.

Convergent validity

Convergent validity was examined by comparing the relationships between the trust measures and other existing scales or variables which have been previously linked with trust. Spearman's rank correlation analyses showed that the trust measures in the current study were significantly correlated with existing related items as expected (Table 5). For example, frequency of buying was linked with product trust – the more one trusts a product, the more likely they are to buy that product. Therefore, convergent validity of each trust factor and the toolkit was established. Convergent validity of the toolkit was further established as none of the trust factors had an AVE of below 0.5 (minimum AVE 0.53 – results not shown).

<Insert Table 5 about here>

Discriminant validity

Initial evidence of discriminant validity was seen in the factor loadings in the CFA in the current study and EFA in study one. None of the items cross-loaded on to more than one factor or type of trust. Table 6 also shows evidence of discriminant validity as the square root of the AVE for each trust factor is greater than the correlation between that factor and the other trust factors. The MSV was also less than the AVE for each factor, further indicating discriminant validity.

<Insert Table 6 about here>

3.3 Discussion

The purpose of study two was to confirm the structure and test the validity of the scales developed in the previous study. Given the emergence of two unexpected distrust factors in study one, the current study began by using confirmatory factor analysis to develop models to test whether these distrust factors may have emerged due to methods effects, as a result of individuals responding differently to these items as they were negatively worded. While

451 the results suggested some improvement in model fit when methods effects were
452 addressed, this was likely as a result of the large number of modifications made to the
453 model. It is therefore suggested that the two distrust factors are true factors and not due to
454 methods effects and patterns of responding. Given the parsimony of the initial confirmed
455 model from study one, this was accepted as the final model.

456 Results of the validity testing suggest that the toolkit has face, convergent, and
457 discriminant validity. Furthermore, the use of four countries in the sample suggests that the
458 toolkit may be applied across different countries and cultures.

459

460 ## 4. Study 3 – Reliability testing

461 Following the development and validation of the factor structure of the toolkit, study 3
462 was designed to assess the reliability of the scales. Both composite and test-retest reliability
463 of the scales were examined.

464

465 ### 4.1 Method

466

467 #### Participants and procedure

468 All participants who fully completed study 2 were re-contacted two weeks after
469 completion to invite participation in study 3. The study took place in November and
470 December 2018. In total, 247 participants were recruited across the four countries (UK n =
471 59, Finland n = 60, Germany n = 58, and Greece n = 70), the mean age was 50.64 (SD =
472 16.35, range = 18 to 65) (Table 7).

473

474 <Insert Table 7 about here>

475

476 Participants completed only the 45 questionnaire items relating to trust. These were the
477 same 45 items included in study 2. Each item used a forced response option, leading to no
478 missing data. The questionnaire took approximately 10 minutes to complete. Ethical
479 approval (10/18/BensonT) was granted by the Queen's University Belfast School of
480 Biological Sciences Research Ethics Committee. The study was conducted in accordance
481 with the Declaration of Helsinki and informed consent was obtained.

482

Each participant's data from study 2 and study 3 were matched using an anonymised code. Where necessary, any reverse scored items were then recoded. Composite (internal) reliability was used to examine agreement between the items in a scale. The composite reliability value for each scale was calculated using the standardised factor loadings of items and their respective error variances. A value of 0.7 or higher shows good reliability (Hair et al., 2014)

The temporal stability of the scales was examined using the Intra-class Correlation Coefficient (ICC). This shows the level of agreement between item answers over a time period (in this case at least two weeks between study 2 and the current study). A stronger ICC indicates greater agreement between answers and therefore suggests greater temporal stability. An ICC value of 0.50-0.75 indicates moderate reliability, a value of 0.75 – 0.90 indicates good reliability, while a value of greater than 0.90 suggests excellent reliability (Koo & Li, 2016). All analyses were conducted using IBM SPSS Statistics v25.

4.2 Results

Table 8 shows the results for the reliability of the toolkit scales. The minimum composite reliability value was 0.82, above the acceptable level of 0.70. In terms of temporal stability, all scales had either moderate (ICC value 0.50 – 0.75) or good (ICC value 0.75 – 0.90) reliability.

<Insert Table 8 about here>

4.3 Discussion

Study three showed that the scales in the trust toolkit were reliable. This was both in terms of internal reliability and a good level of agreement between the items in each scale, and temporal stability with a good level of agreement between each participant's scores over time.

5. General discussion

Consumer trust is an important aspect of the food market. Consumers expect that goods purchased are authentic, safe, and of satisfactory quality. Food safety incidents and

515 modernisation in the food supply chain have arguably not only led to decreases in consumer

516 trust, but have also simultaneously led to a growing importance in trust as consumers

517 become further distanced from development and production. In order to understand if

518 attempts to improve consumer trust are successful, it is necessary to have a valid and

519 reliable method of measuring trust. While a plethora of studies have measured different

520 aspects of consumer trust in relation to food, many of the measures developed lack

521 validation and/or reliability. In order to ensure that changes following efforts to increase

522 consumer trust represent true effects, validation and reliability of scales are vital. The aim of

523 the current set of studies was to develop and test a consumer trust toolkit consisting of items

524 which can be used to measure consumer trust in relation to various actors or aspects of the

525 food system. The results suggest that the scales developed contain accurate items which

526 may be used to measure different aspects of consumer trust.

527   A review of the literature suggested that there were five different types of trust

528 relating to the food chain: Interpersonal trust, general organisation trust, specific organisation

529 trust, food chain trust, and product trust. However, results from our exploratory factor

530 analysis suggested six different factors: Interpersonal trust, organisation trust, food chain

531 trust, product trust, interpersonal distrust, and organisation distrust. Inspection of the factors

532 showed that those items we believed measured specific trust in organisations grouped

533 together with those items we believed measured (general) organisation trust. This suggests

534 that consumers do not distinguish between general trust in organisations and trust in

535 organisations to perform certain tasks. That is, if an individual trusts an organisation in

536 general, then this trust appears to extend to trust in their ability to perform any specific tasks.

537 In the literature, trust is separated into general versus specific trust, where general trust is

538 referred to as interpersonal trust or a personality trait and specific trust as trust in a specific

539 entity or object (Stefani, Cavicchi, Romano, & Lobb, 2008). Referring back to the definition of

540 trust provided in the introduction as "the willingness of a party to be vulnerable to the actions

541 of another party based on the expectation that the other will perform a particular action

542 important to the trustor, irrespective of the ability to monitor or control that other party"

543 (Mayer et al., 1995), it is logical that specific trust relates to a specific entity rather than a

544 specific task. This indistinguishable link between general trust in an organisation and trust in

545 the organisation's ability to perform certain tasks means that entities involved in food

546 production and supply have some flexibility with regards to their actions, as these are trusted

547 by the consumer. When an organisation's tasks or actions are considered to be honest and

548 sincere, this may produce a halo effect of improving the general trust in that organisation.

549   A further interesting finding from the results was the identification of two distrust

550 factors – interpersonal distrust and organisation distrust. This was unexpected given that we

551 did not believe that any items in the questionnaire measured distrust. However, several
552 items measuring trust were negatively worded, that is, they differed in their wording
553 compared to most other items in the questionnaire. An example of a positively worded item
554 was "most people are trustworthy" while a negatively worded item was "you can't trust
555 strangers anymore". The inclusion of negatively worded items is designed to reduce
556 acquiescence error or bias, when a participant answers affirmatively to all items regardless
557 of content (Hinz, Michalski, Schwarz, & Herzberg, 2007). For example, if a participant
558 answers '7' on a scale of 1 to 7 on both all positive and negative items, one may infer the
559 participant did not attend to or understand the items as these assess very different opinions.
560 With regards to questionnaire development, recent research suggests that the inclusion of
561 negative items may lead to the formation of a methods factor based on how participants
562 respond rather than a 'true' factor (X. Zhang et al., 2016). In study 2, the use of CFA allowed
563 for the examination of potential methods effects. While the models accounting for methods
564 effects showed marginal improvements in model fit over the standard model, this may have
565 been due to the large number of modifications made to the revised models to account for
566 potential methods effects. That the negative items loaded on to two factors as opposed to
567 one factor suggests that methods effects may not be the only explanation. The body of
568 evidence recognising that trust and distrust are related yet distinct concepts (Cho, 2006; Lee
569 et al., 2018; Lewicki, McAllister, & Bies, 1998; McKnight, Kacmar, & Choudhury, 2004) also
570 supports our contention that the distrust factors in the current studies emerged due to a
571 conceptual difference between trust and distrust rather than methods effects. However,
572 given the unequal and limited number of negatively worded items compared to positive
573 items, as well as the use of negatively worded items for only some types of trust, we were
574 unable to investigate this fully. Future research might investigate this issue further for the
575 different types of trust and using an equal number of positively and negatively worded items.
576 Future studies in this area should be aware of the implications of using negatively worded
577 trust items or reverse wording trust items. This may lead to measuring the different concept
578 of distrust rather than trust. Furthermore, the inclusion of both positively and negatively
579 worded items may cause additional issues with regards to respondent confusion and
580 consistency (Colosi, 2005; Salazar, 2015).

581 The final factor model details a rational and logical solution supported by the data. All
582 items load effectively on one factor only and the relationships between factors shows distinct
583 but related concepts. Relationships between the factors are as would be expected. For
584 example, product trust has a stronger relationship with chain trust and organisation trust than
585 the other types of trust. Multiple different types of validity tests were conducted and the use
586 of different methods to assess each type of validity as well as reliability testing is a particular

587   strength of the current set of studies. The sampling and testing of multiple countries with
588   varying levels of trust according to previous research means that the toolkit has broad
589   application. The use of a relatively large sample size for factor analysis and testing was also
590   a strength.

591   A limitation of the current study was the use of an EU only sample. As such, in
592   addition to testing for further methods effects, future studies should sample other countries.
593   This is particularly pertinent given that trust can vary widely between countries in the East
594   and West (Krockow, Takezawa, Pulford, Colman, & Kita, 2017; Yamagishi & Yamagishi,
595   1994). In countries outside the EU, information on food and safety or quality may be
596   provided by those involved in the chain such as manufacturers and retailers rather than an
597   independent organisation related to the chain such as EFSA. While the toolkit measures
598   trust in those involved in the chain, it does not specifically assess trust in information
599   provided by those in the chain. The toolkit was created with the aim of being adaptable to
600   various specific aspects of the food system such as different products and actors within the
601   chain. While we used specific examples such as EFSA, food manufacturers, and beef
602   burgers in the items used in the studies, we created the toolkit to be adaptable and believe
603   the items in the toolkit can be applied to different specific aspects of the food system such as
604   different products and actors. Future research should examine how the validity and reliability
605   of the toolkit is affected by using different specifics. Finally, the accuracy of the toolkit might
606   also be tested in further studies by using the toolkit to measure baseline trust, intervening to
607   increase trust then measuring again using the toolkit to see if there has been a resulting
608   increase in trust.

609

6. Conclusions

611   The consumer trust toolkit is a valid and reliable collection of items to measure trust
612   in the food system. Drawing upon previous research, the toolkit contains items to measure
613   trust in various levels of the food system from production through to consumption. Given the
614   relatively modular nature of the toolkit, researchers in this area can use a specific collection
615   of items to measure trust depending upon which aspects they are most interested.
616   Consumer trust in food is currently low and the toolkit can be used in future studies to
617   identify the most effective methods to improve trust.

618

619

629 ## Declarations of interest

630 None

631

632 **References**

633 Adams, J. E., Highhouse, S., & Zickar, M. J. (2010). Understanding general distrust
634   of corporations. *Corporate Reputation Review*, *13*(1), 38–51. JOUR.
635   http://doi.org/10.1057/crr.2010.6

636 Allum, N. (2007). An empirical test of competing theories of hazard related trust:
637   The case of GM food. *Risk Analysis: An International Journal*, *27*(4), 935–946.
638   JOUR.

639 Ariyawardana, A., Ganegodage, K., & Mortlock, M. Y. (2017). Consumers' trust in
640   vegetable supply chain members and their behavioural responses: A study
641   based in Queensland, Australia. *Food Control*, *73*, 193–201. JOUR.

642 Berg, L. (2004). Trust in food in the age of mad cow disease: a comparative study of
643   consumers' evaluation of food safety in Belgium, Britain and Norway. *Appetite*,
644   *42*(1), 21–32.

645 Bozic, B. (2017). Consumer trust repair: A critical literature review. *European*
646   *Management Journal*, *35*(4), 538–547.

647 Brudvig, S. (2015). Consumer-based brand trust scales: validation and assessment.
648   In *Revolution in Marketing: Market Driving Changes* (pp. 17–21). CHAP,
649   Springer.

650 Bruschi, V., Shershneva, K., Dolgopolova, I., Canavari, M., & Teuber, R. (2015).
651   Consumer perception of organic food in emerging markets: evidence from Saint
652   Petersburg, Russia. *Agribusiness*, *31*(3), 414–432.

653 Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for
654   researchers. *Communication Methods and Measures*, *12*(1), 25–44.

655 Cho, J. (2006). The mechanism of trust and distrust formation and their relational
656   outcomes. *Journal of Retailing*, *82*(1), 25–35.
657   http://doi.org/10.1016/J.JRETAI.2005.11.002

658 Colosi, R. (2005). Negatively Worded Questions Cause Respondent Confusion. In
659   *Annual Meeting American Statistical Association* (pp. 2896–2903). Minneapolis,
660   MN. Retrieved from
661   http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.556.243&rep=rep1&ty
662   pe=pdf

663 Connolly, R., & Bannister, F. (2007). Consumer trust in Internet shopping in Ireland:
664   towards the development of a more effective trust measurement instrument.
665   *Journal of Information Technology*, *22*(2), 102–118.

666 Coveney, J., Henderson, J., Mayer, S., Mamerow, L., Taylor, A., Ward, P., & Wilson,
667   A. (2015). Consumer Trust. In K. Albala (Ed.), *The SAGE Encyclopedia of Food*
668   *Issues* (First, pp. 282–285). Thousand Oaks, CA, CA: Sage Publications Sage
669   CA: Thousand Oaks, CA.

670 De Jonge, J., Van Trijp, H., Jan Renes, R., & Frewer, L. (2007). Understanding
671   consumer confidence in the safety of food: Its two dimensional structure and
672   determinants. *Risk Analysis*, *27*(3), 729–740. JOUR.

673 de Jonge, J., van Trijp, J. C. M., van der Lans, I. A., Renes, R. J., & Frewer, L. J.
674   (2008). How trust in institutions and organizations builds general consumer
675   confidence in the safety of food: A decomposition of effects. *Appetite*, *51*(2),
676   311–317. http://doi.org/10.1016/J.APPET.2008.03.008

677 Ding, Y., Veeman, M. M., & Adamowicz, W. L. (2012). The impact of generalized
678   trust and trust in the food system on choices of a functional GM food.
679   *Agribusiness*, *28*(1), 54–66.

680 Dumortier, J., Evans, K. S., Grebitus, C., & Martin, P. A. (2017). The influence of
681   trust and attitudes on the purchase frequency of organic produce. *Journal of*

682   *International Food & Agribusiness Marketing*, *29*(1), 46–69.

683   Edelman. (2018). *2018 Edelman Trust Barometer: Global Report*. Retrieved from
684   https://www.edelman.com/sites/g/files/aatuss191/files/2018-
685   10/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf

686   Etienne, J., Chirico, S., McEntaggart, K., Papoutsis, S., & Millstone, E. (2018). EU
687   Insights – Consumer perceptions of emerging risks in the food chain. *EFSA*
688   *Supporting Publications*, *15*(4), 1394E. JOUR.
689   http://doi.org/10.2903/sp.efsa.2018.EN-1394

690   European Commission. (2012). *Special Eurobarometer 389: EUROPEANS'*
691   *ATTITUDES TOWARDS FOOD SECURITY, FOOD QUALITY AND THE*
692   *COUNTRYSIDE*. Retrieved from
693   http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_389_en.pdf

694   European Food Safety Authority. (2010). *SPECIAL EUROBAROMETER 354: Food-*
695   *related risks*. Parma: Italy. Retrieved from
696   http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_354_en.pdf

697   European Social Survey. (2012). *ESS Round 6 Source Questionnaire*. London.
698   Retrieved from
699   https://www.europeansocialsurvey.org/docs/round6/fieldwork/source/ESS6_sour
700   ce_main_questionnaire.pdf

701   Field, A. (2009). *Discovering Statistics Using SPSS* (Second). London, UK: Sage
702   Publications.

703   Frewer, L. J., & Miles, S. (2003). Temporal stability of the psychological determinants
704   of trust: Implications for communication about food risks. *Health, Risk & Society*,
705   *5*(3), 259–271.

706   Frewer, L. J., Scholderer, J., & Bredahl, L. (2003). Communicating about the risks
707   and benefits of genetically modified foods: The mediating role of trust. *Risk*
708   *Analysis: An International Journal*, *23*(6), 1117–1133. JOUR.

709   Fritz, M., Martino, G., & Surci, G. (2008). Trust conditional on governance structure:
710   theory and evidence from case studies. *Journal on Chain and Network Science*,
711   *8*(1), 33–46.

712   Giampietri, E., Verneau, F., Del Giudice, T., Carfora, V., & Finco, A. (2018). A
713   Theory of Planned behaviour perspective for investigating the role of trust in
714   consumer purchasing decision related to short food supply chains. *Food Quality*
715   *and Preference*, *64*, 160–166.

716   Gurviez, P., & Korchia, M. (2003). Proposal for a multidimensional brand trust scale.
717   In *32nd Emac-Conference-Glasgow, Marketing: Responsible and Relevant* (pp.
718   438–452). CONF.

719   Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data*
720   *Analysis* (Seventh). Essex: Pearson Education.

721   Hinkin, T. R., Tracey, J. B., & Enz, C. A. (1997). Scale Construction: Developing
722   Reliable and Valid Measurement Instruments. *Journal of Hospitality & Tourism*
723   *Research*, *21*(1), 100–120. http://doi.org/10.1177/109634809702100108

724   Hinz, A., Michalski, D., Schwarz, R., & Herzberg, P. Y. (2007). The acquiescence
725   effect in responding to a questionnaire. *Psycho-Social Medicine*, *4*, Doc07.
726   Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19742288

727   Hobbs, J. (2011). Recovering consumer confidence after a food crisis. *Advances in*
728   *Pork Production*, *22*, 217–224.

729   Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling:
730   Guidelines for determining model fit. *Articles*, 2.

731   Jin, S., & Zhou, L. (2014). Consumer interest in information provided by food

traceability systems in Japan. *Food Quality and Preference*.
http://doi.org/10.1016/j.foodqual.2014.04.005

Jokinen, P., Kupsala, S., & Vinnari, M. (2012). Consumer trust in animal farming practices - exploring the high trust of Finnish consumers. *International Journal of Consumer Studies*, *36*(1), 106–113. http://doi.org/10.1111/j.1470-6431.2011.00996.x

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36.

Kendall, H., Naughton, P., Kuznesof, S., Raley, M., Dean, M., Clark, B., … Zhong, Q. (2018). Food fraud and the perceived integrity of European food imports into China. *PloS One*, *13*(5), e0195817. JOUR.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.

Krockow, E. M., Takezawa, M., Pulford, B. D., Colman, A. M., & Kita, T. (2017). Cooperation and trust in Japanese and British samples: Evidence from incomplete information games. *International Perspectives in Psychology: Research, Practice, Consultation*, *6*(4), 227–245. http://doi.org/10.1037/ipp0000074

Lassoued, R., Hobbs, J. E., Micheels, E. T., & Zhang, D. Di. (2015). Consumer Trust in Chicken Brands: A Structural Equation Model. *Canadian Journal of Agricultural Economics/Revue Canadienne d'agroeconomie*, *63*(4), 621–647. http://doi.org/10.1111/cjag.12082

Lee, S.-J., Ahn, C., Song, K. M., Ahn, H., Lee, S.-J., Ahn, C., … Ahn, H. (2018). Trust and Distrust in E-Commerce. *Sustainability*, *10*(4), 1015. http://doi.org/10.3390/su10041015

Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, *23*(3), 438–458.

Liu, R., Gao, Z., Nayga Jr, R. M., Snell, H. A., & Ma, H. (2019). Consumers' valuation for food traceability in China: Does trust matter? *Food Policy*, *88*, 101768.

Lu, J., Wu, L., Wang, S., & Xu, L. (2016). Consumer preference and demand for traceable food attributes. *British Food Journal*, *118*(9), 2140–2156.

MacLeod, H. (2013). Views on the horsemeat scandal. Retrieved December 3, 2018, from https://yougov.co.uk/topics/consumer/articles-reports/2013/02/15/views-horsemeat-scandal

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734.

McKnight, D. H., Kacmar, C. J., & Choudhury, V. (2004). Dispositional trust and distrust distinctions in predicting high-and low-risk internet expert advice site perceptions. *E-Service*, *3*(2), 35–58.

Meixner, O., Ameseder, C., Haas, R., Canavari, M., Fritz, M., & Hofstede, G. J. (2009). Importance of trust building elements in business-to-business agri-food chains. *Journal of Farm Management*, *13*(9), 655–668.

Menozzi, D., Halawany-Darson, R., Mora, C., & Giraud, G. (2015). Motives towards traceable food choice: A comparison between French and Italian consumers. *Food Control*, *49*, 40–48.

Miran, B., & Akgüngör, S. (2005). The effect of mad cow (BSE) scare on beef demand and sales loss: The case of Izmir. *Turkish Journal of Veterinary and*

782   *Animal Sciences*, *29*(2), 225–231.

783   Moorman, C., Deshpande, R., & Zaltman, G. (1993). Factors affecting trust in market
784   research relationships. *Journal of Marketing*, *57*(1), 81–101.

785   Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship
786   marketing. *Journal of Marketing*, *58*(3), 20–38.

787   Nuttavuthisit, K., & Thøgersen, J. (2017). The importance of consumer trust for the
788   emergence of a market for green products: The case of organic food. *Journal of*
789   *Business Ethics*, *140*(2), 323–337. JOUR.

790   Office for National Statistics. (2015). Measuring National Well-being - Office for
791   National Statistics. Retrieved April 29, 2019, from
792   https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/bulletins/mea
793   suringnationalwellbeing/2015-09-23

794   Perrini, F., Castaldo, S., Misani, N., & Tencati, A. (2010). The impact of corporate
795   social responsibility associations on trust in organic products marketed by
796   mainstream retailers: a study of Italian consumers. *Business Strategy and the*
797   *Environment*, *19*(8), 512–526. JOUR.

798   PriceWaterhouseCoopers. (2015). *Food trust - Giving customers confidence in your*
799   *food*. Delaware, Belfast. Retrieved from
800   https://www.pwc.nl/nl/assets/documents/pwc-food-supply-and-integrity-
801   services.pdf

802   Rebelo-Pinto, T., Pinto, J. C., Rebelo-Pinto, H., & Paiva, T. (2014). Validation of a
803   three-dimensional model about sleep: Habits, personal factors and
804   environmental factors. *Sleep Science*, *7*(4), 197–202.
805   http://doi.org/10.1016/J.SLSCI.2014.12.002

806   Roininen, K., Lähteenmäki, L., & Tuorila, H. (1999). Quantification of consumer
807   attitudes to health and hedonic characteristics of foods. *Appetite*, *33*(1), 71–88.

808   Roosen, J., Lusk, J. L., & Fox, J. A. (2003). Consumer demand for and attitudes
809   toward alternative beef labeling strategies in France, Germany, and the UK.
810   *Agribusiness*, *19*(1), 77–90. http://doi.org/10.1002/agr.10041

811   Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust 1.
812   *Journal of Personality*, *35*(4), 651–665.

813   Salazar, M. S. (2015). The dilemma of combining positive and negative items in
814   scales. *Psicothema*, *27*(2), 192–199.

815   Schlenker, W., & Villas-Boas, S. B. (2009). Consumer and market responses to mad
816   cow disease. *American Journal of Agricultural Economics*, *91*(4), 1140–1152.

817   Siegrist, M., Earle, T. C., & Gutscher, H. (2003). Test of a trust and confidence
818   model in the applied context of electromagnetic field (EMF) risks. *Risk Analysis:*
819   *An International Journal*, *23*(4), 705–716. JOUR.

820   Spence, M., Stancu, V., Dean, M., Livingstone, M. B. E., Gibney, E. R., &
821   Lähteenmäki, L. (2016). Are food-related perceptions associated with meal
822   portion size decisions? A cross-sectional study. *Appetite*, *103*, 377–385.

823   Stefani, G., Cavicchi, A., Romano, D., & Lobb, A. E. (2008). Determinants of
824   intention to purchase chicken in Italy: the role of consumer risk perception and
825   trust in different information sources. *Agribusiness*, *24*(4), 523–537.
826   http://doi.org/10.1002/agr.20177

827   Tonkin, E., Wilson, A. M., Coveney, J., Meyer, S. B., Henderson, J., McCullum, D.,
828   … Ward, P. R. (2019). Consumers respond to a model for (re)building consumer
829   trust in the food system. *Food Control*, *101*, 112–120.
830   http://doi.org/10.1016/J.FOODCONT.2019.02.012

831   Wilson, A. M., Withall, E., Coveney, J., Meyer, S. B., Henderson, J., McCullum, D.,

832 … Ward, P. R. (2016). A model for (re) building consumer trust in the food
833 system. *Health Promotion International*, *32*(6), 988–1000.
834 Worthington, R. L., & Whittaker, T. A. (2006). Scale Development Research: A
835 Content Analysis and Recommendations for Best Practices. *The Counseling*
836 *Psychologist*, *34*(6), 806–838. http://doi.org/10.1177/0011000006288127
837 Wu, L., Wang, H., Zhu, D., Hu, W., & Wang, S. (2016). Chinese consumers'
838 willingness to pay for pork traceability information—The case of Wuxi.
839 *Agricultural Economics*, *47*(1), 71–79.
840 Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States
841 and Japan. *Motivation and Emotion*, *18*(2), 129–166. JOUR.
842 http://doi.org/10.1007/BF02249397
843 Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on
844 exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*,
845 *9*(2), 79–94.
846 Zachmann, K., & Østby, P. (2011). Food, technology, and trust: an introduction.
847 *History and Technology*, *27*(1), 1–10.
848 http://doi.org/10.1080/07341512.2011.548970
849 Zhang, X., Noor, R., & Savalei, V. (2016). Examining the Effect of Reverse Worded
850 Items on the Factor Structure of the Need for Cognition Scale. *PLOS ONE*,
851 *11*(6), e0157795. http://doi.org/10.1371/journal.pone.0157795
852 Zhang, Y., Jing, L., Bai, Q., Shao, W., Feng, Y., Yin, S., & Zhang, M. (2018).
853 Application of an integrated framework to examine Chinese consumers'
854 purchase intention toward genetically modified food. *Food Quality and*
855 *Preference*, *65*, 118–128. JOUR.
856
857

858

859
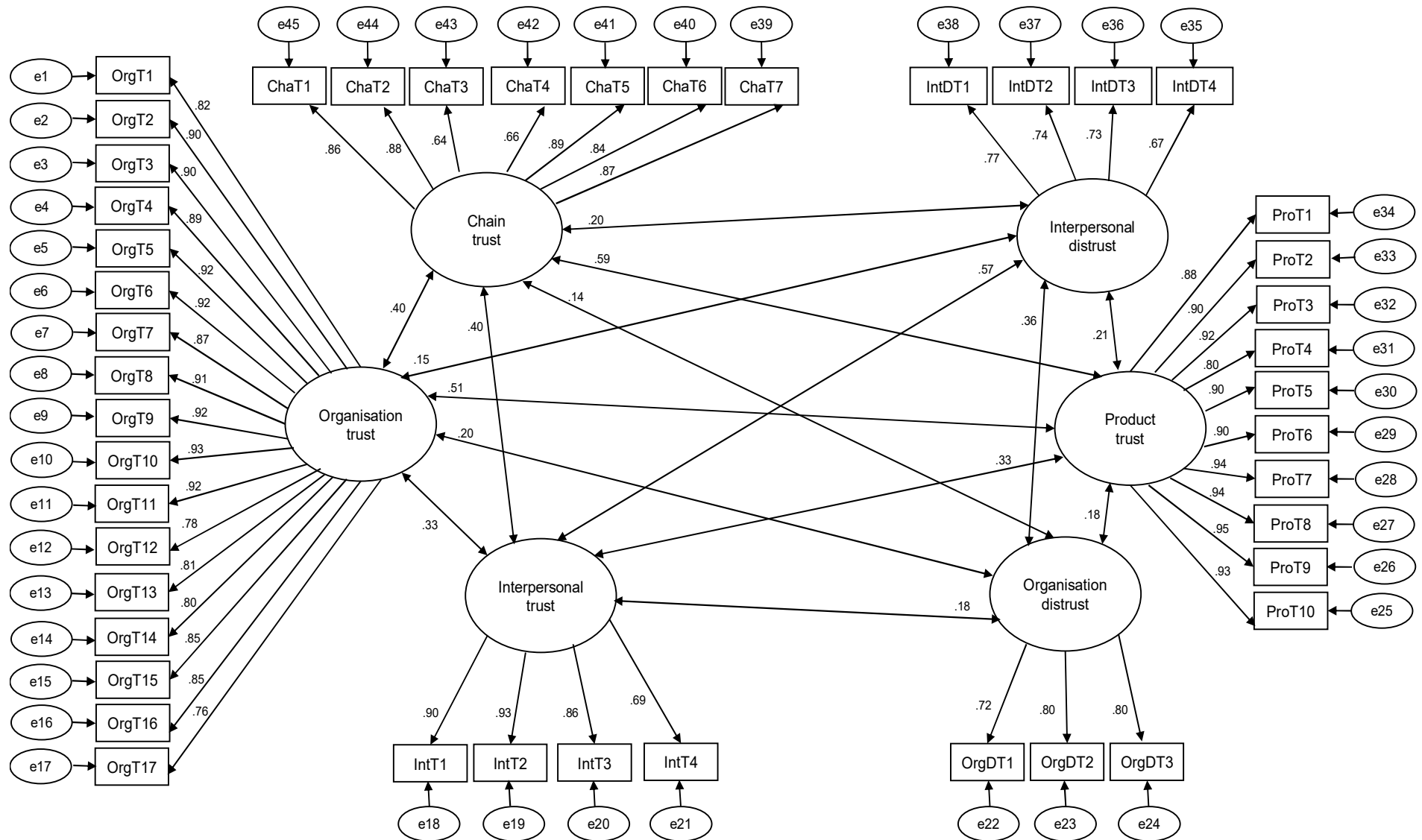
Figure 1: Final measurement model for consumer trust toolkit with standardised factor loadings and correlations

861 **Table 1: Characteristics of participants in study 1 conducted in the UK**

Characteristic	%
	100%
Gender	
Male	49%
Female	51%
Age	
18-24	13%
25-34	17%
35-44	18%
45-54	18%
55-64	15%
65+	20%
Highest level of completed education	
Primary school only or incomplete secondary education	4%
Completed secondary education (GCSE)	19%
A-Level or vocational qualification	37%
Undergraduate degree	28%
Postgraduate degree or doctorate	11%
Prefer not to answer	1%
Marital status	
Married or living with partner	65%
Never married	25%
Separated/widowed/divorced	10%
Prefer not to answer	1%

862   * Percentages may total more than 100% due to rounding

863

864

865

866 **Table 2: Characteristics of participants for each country in study 2**

Characteristic/country	UK	Germany	Greece	Finland	Total n (%)
	100%	100%	100%	100%	100%
Gender					
Male	49%	46%	46%	47%	49%
Female	50%	54%	54%	53%	51%
Other	0%	0%	0%	0%	0%
Age					
18-24	12%	11%	12%	13%	12%
25-34	16%	15%	22%	15%	17%
35-44	14%	16%	21%	18%	17%
45-54	18%	18%	18%	15%	17%
55-64	15%	14%	16%	10%	14%
65+	25%	27%	12%	28%	23%
Highest level of completed education					
Primary school only or incomplete secondary education	6%	6%	1%	11%	6%
Completed secondary education (GCSE)	16%	17%	18%	41%	23%
A-Level or vocational qualification	39%	55%	20%	21%	34%
Undergraduate degree	25%	6%	29%	11%	20%
Postgraduate degree or doctorate	14%	16%	22%	12%	16%
Prefer not to answer	0%	0%	1%	4%	1%
Marital status					
Married or living with partner	58%	60%	62%	50%	57%
Never married	25%	22%	25%	30%	26%
Separated/widowed/divorced	16%	17%	11%	7%	15%
Prefer not to answer	1%	1%	2%	2%	2%

867 * Percentages may add to more than 100% due to rounding

868

869

870 **Table 3: Fit statistics for each CFA model**

	χ^2 (p), χ^2/df	RMSEA	CFI	NFI	TLI
Model 1	6371.67 *(p < 0.001),* 6.9	0.08	0.90	0.88	0.89
Model 2	6423.73 *(p < 0.001),* 6.9	0.08	0.90	0.88	0.89
Model 3	4733.20 *(p < 0.001),* 5.8	0.07	0.93	0.91	0.91

871   χ^2, chi-square; RMSEA, Root Mean Square Error of Approximation; CFI,

872   Comparative Fit Index; NFI, Normed-Fit Index; TLI, Tucker-Lewis Index

Table 4: Standardised factor loadings for each item by factor for final accepted model (model 1)

Label	Item	Factor	Organisation trust	Product trust	Interpersonal trust	Chain trust	Organisation distrust	Interpersonal distrust
OrgT1	You can count on EFSA[1]		0.82	-	-	-	-	-
OrgT2	I trust EFSA[1]		0.90	-	-	-	-	-
OrgT3	Consumers can always rely on EFSA[1]		0.90	-	-	-	-	-
OrgT4	EFSA keep their promises[1]		0.89	-	-	-	-	-
OrgT5	I believe in EFSA[2]		0.92	-	-	-	-	-
OrgT6	I have confidence in EFSA[2]		0.92	-	-	-	-	-
OrgT7	EFSA make me feel safe[3]		0.87	-	-	-	-	-
OrgT8	EFSA is sincere with consumers[3]		0.91	-	-	-	-	-
OrgT9	EFSA is honest with consumers[3]		0.92	-	-	-	-	-
OrgT10	EFSA is dependable[4]		0.93	-	-	-	-	-
OrgT11	I trust EFSA to provide accurate information[5]		0.92	-	-	-	-	-
OrgT12	EFSA has a good understanding of all the issues relevant[6]		0.78	-	-	-	-	-
OrgT13	EFSA take their responsibility to society seriously[6]		0.81	-	-	-	-	-
OrgT14	EFSA are good at looking at the evidence and judging what to do[6]		0.80	-	-	-	-	-
OrgT15	EFSA has practices that favour the consumer's best interests[4]		0.85	-	-	-	-	-
OrgT16	EFSA considers the consumer's welfare when making important decisions[4]		0.85	-	-	-	-	-
OrgT17	EFSA considers how future decisions and actions will affect the consumer[4]		0.76	-	-	-	-	-
ProT1	I trust that EU beef burgers are high quality[7]		-	0.88	-	-	-	-
ProT2	EU beef burgers are reliable[7]		-	0.90	-	-	-	-
ProT3	I trust that EU beef burgers are safe[8]		-	0.92	-	-	-	-
ProT4	I trust that EU beef burgers are fully traceable back to their origin[9]		-	0.80	-	-	-	-
ProT5	I trust that EU beef burgers are authentic[9]		-	0.90	-	-	-	-
ProT6	I trust that EU beef burgers are accurately labelled		-	0.90	-	-	-	-
ProT7	EU beef burgers are trustworthy		-	0.94	-	-	-	-
ProT8	EU beef burgers are honest		-	0.94	-	-	-	-
ProT9	EU beef burgers are truthful		-	0.95	-	-	-	-
ProT10	EU beef burgers have integrity		-	0.93	-	-	-	-

Label	Item	Factor	Organisation trust	Product trust	Interpersonal trust	Chain trust	Organisation distrust	Interpersonal distrust
IntT1	Most people are basically honest[10]		-	-	0.90	-	-	-
IntT2	Most people are trustworthy[10]		-	-	0.93	-	-	-
IntT3	Most people are basically good and kind[10]		-	-	0.86	-	-	-
IntT4	Most people are trustful of others[10]		-	-	0.69	-	-	-
ChaT1	Food manufacturers take good care of the safety of our food[11]		-	-	-	0.86	-	-
ChaT2	Food manufacturers give special attention to the safety of food[11]		-	-	-	0.88	-	-
ChaT3	Food manufacturers have the competence to control the safety of food[11]		-	-	-	0.64	-	-
ChaT4	Food manufacturers have sufficient knowledge to guarantee the safety of food products[11]		-	-	-	0.66	-	-
ChaT5	Food manufacturers are honest about the safety of food[11]		-	-	-	0.89	-	-
ChaT6	Food manufacturers are sufficiently open regarding the safety of food[11]		-	-	-	0.84	-	-
ChaT7	Food manufacturers can be trusted to protect the consumer from unsafe food[12]		-	-	-	0.87	-	-
OrgDT1	Information from EFSA is distorted[13]		-	-	-	-	0.72	-
OrgDT2	Information from EFSA has been proven wrong in the past[13]		-	-	-	-	0.80	-
OrgDT3	EFSA provides accurate information only to protect themselves and their own interests[13]		-	-	-	-	0.80	-
IntDT1	If given a chance, most people would try to take advantage of you[14]		-	-	-	-	-	0.77
IntDT2	Most people are too busy looking out for themselves to be helpful[14]		-	-	-	-	-	0.74
IntDT3	You can't trust strangers anymore[14]		-	-	-	-	-	0.73
IntDT4	I never rely on other people[14]		-	-	-	-	-	0.67

874   [1] Item adapted from Perrini et al. (2010)
875   [2] Item adapted from Nuttavuthisit & Thøgersen (2017)
876   [3] Item adapted from Gurviez & Korchia (2003)
877   [4] Item adapted from Brudvig (2015)
878   [5] Item adapted from Zhang et al. (2018)
879   [6] Item adapted from Allum (2007)
880   [7] Item adapted from Lassoued et al. (2015)
881   [8] Item adapted from Ariyawardana, Ganegodage, & Mortlock (2017)

882   [9] Item adapted from Spence et al. (2016)
883   [10] Item from Yamagishi & Yamagishi (1994)
884   [11] Item adapted from de Jonge, van Trijp, van der Lans, Renes, & Frewer (2008)
885   [12] Item adapted from Kendall et al. (2018)
886   [13] Item adapted from Frewer, Scholderer, & Bredahl (2003)
887   [14] Item from Siegrist, Earle & Gutscher (2003)

888 **Table 5: Associations between existing scales and consumer toolkit scales for**
889 **convergent validity**

Consumer toolkit scale	Previously associated or existing scale	r_s
Interpersonal trust	Commonly used general trust item (1 item)	0.64**
Interpersonal trust	Life satisfaction	0.31**
Organisation trust	Number of sources of information wanted if risk to food safety	0.13**
Organisation trust	Corporate distrust scale	- 0.40*
Chain trust	Frequency of buying	0.16**
Chain trust	Food quality interest	- 0.11**
Product trust	Frequency of buying	0.41**
Product trust	Food quality interest	- 0.17**

890 ** $p < 0.01$

891

892

893

894 **Table 6: Correlations, square root of Average Variance Extracted (AVE) and Maximum**
895 **Shared Variance (MSV) for each of the scales in the toolkit**

Scale	IT	ID	OT	OD	CT	PT	MSV
Interpersonal Trust (IT)	**0.85**	-	-	-	-	-	0.32
Interpersonal Distrust (ID)	0.57	**0.73**	-	-	-	-	0.32
Organisation Trust (OT)	0.33	0.15	**0.87**	-	-	-	0.26
Organisation Distrust (OD)	0.18	0.36	0.20	**0.77**	-	-	0.13
Chain Trust (CT)	0.40	0.20	0.46	0.14	**0.81**	-	0.35
Product Trust (PT)	0.33	0.21	0.51	0.18	0.59	**0.91**	0.35

896 *Square root of AVE shown in bold on diagonal*
897 *Note: The trust and distrust scales are positively correlated as the distrust items have been*
898 *reverse scored*

899

900

901

902   **Table 7: Characteristics of participants for each country in study 3**

Characteristic/country	UK	Germany	Greece	Finland	Total n (%)
	100%	100%	100%	100%	100%
Gender					
Male	56%	43%	46%	45%	47%
Female	44%	57%	54%	55%	53%
Age					
18-24	2%	5%	11%	7%	7%
25-34	10%	10%	26%	8%	14%
35-44	17%	19%	16%	23%	19%
45-54	20%	19%	13%	18%	17%
55-64	14%	10%	20%	7%	13%
65+	37%	36%	14%	37%	30%
Highest level of completed education					
Primary school only or incomplete secondary education	9%	7%	0%	10%	6%
Completed secondary education (GCSE)	20%	19%	21%	42%	26%
A-Level or vocational qualification	37%	59%	19%	20%	33%
Undergraduate degree	20%	5%	30%	13%	18%
Postgraduate degree or doctorate	14%	10%	29%	12%	17%
Prefer not to answer	0%	0%	1%	3%	1%
Marital status					
Married or living with partner	58%	64%	61%	47%	58%
Never married	19%	17%	26%	32%	24%
Separated/widowed/divorced	24%	19%	11%	20%	18%
Prefer not to answer	0%	0%	1%	2%	1%

903   * Percentages may add to more than 100% due to rounding

904

905

906

907 **Table 8: Composite reliability (CR) and Intra-class correlation coefficient (ICC) values**
908 **for each of the scales in the toolkit**

Consumer toolkit scale	Composite reliability	Intra-class correlation
Interpersonal trust	0.91	0.72
Interpersonal distrust	0.82	0.79
Organisation trust	0.98	0.68
Organisation distrust	0.82	0.53
Chain trust	0.93	0.72
Product trust	0.98	0.69

909

910