

Full title: The development and validation of a survey to measure fecal-oral child exposure to zoonotic enteropathogens: The FECEZ Enteropathogens Index

Short title: FECEZ Enteropathogens Index to measure child exposure to zoonotic enteropathogens

April M. Ballard^{a,b}, Regine Haardörfer^c, Betty Corozo Angulo^d, Matthew C. Freeman^{b,h}, Joseph N.S. Eisenberg^e, Gwenyth O. Lee^f, Karen Levy^g, Bethany A. Caruso^{b,c,h}

^a Department of Population Health Sciences, Georgia State University School of Public Health

^b Gangarosa Department of Environmental Health, Emory University Rollins School of Public Health

^c Department of Behavioral, Social, and Health Education Sciences, Emory University Rollins School of Public Health

^d Universidad Técnica Luis Vargas Torres de Esmeraldas

^e Department of Epidemiology, University of Michigan School of Public Health

^f Rutgers Global Health Institute and Department of Biostatistics and Epidemiology, Rutgers School of Public Health

^g Department of Environmental and Occupational Health Sciences, University of Washington School of Public Health

^h Hubert Department of Global Health, Emory University Rollins School of Public Health

Corresponding author

Email: klevyx@uw.edu

1 **Abstract**

2 Child exposure to animal feces and associated enteropathogens contribute to a significant burden of
3 disease in low- and middle-income countries. However, there are no standardized, validated survey-based
4 approaches to enable accurate assessment of child exposure to zoonotic enteropathogens. We developed
5 and validated a survey-based measure of fecal-oral child exposure to zoonotic enteropathogens, the
6 FECEZ Enteropathogens Index. First, we identified critical attributes of child exposure through in-depth
7 interviews (IDIs) in Ecuador among individuals who care for animals ($n=29$) and mothers of children
8 under two years old ($n=58$), and through a systematic review of existing exposure measures. Second,
9 based on these findings, we developed a 105-question survey and administered it to 297 mothers with
10 children under age five. Third, we refined the survey, using principal component analysis to determine the
11 optimal number of components. The final index consisted of 34 items across two sub-domains: the child
12 *Environment* and child *Behavior*. Lastly, we compared index scores to two commonly used, unvalidated
13 measures of child exposure – maternal reported household animal ownership and presence of animal
14 feces. Using the FECEZ Enteropathogens Index revealed varying degrees of exposure in our study
15 population, with only two children having no exposure. In contrast, if we had used animal ownership or
16 the presence of animal feces as a measure of exposure, 44% and 33% of children would have been
17 classified as having no exposure, respectively. These common binary exposure measures may be
18 inadequate because they do not provide sufficient information to identify the relative risk of zoonotic
19 pathogen exposure. The FECEZ Enteropathogens Index overcomes this limitation, advancing our ability
20 to assess exposure by quantifying the multiple components of child exposure to zoonotic enteropathogens
21 with higher resolution. Additional testing and evaluation of the index is needed to ensure its reliability,
22 validity, and cross-cultural equivalence in other contexts.

23 Introduction

24 Enteric pathogens pose serious health risks for children under age five. In addition to acute diarrheal
25 disease (the fifth leading cause of death among children under five years), persistent exposure and
26 recurrent enteric infections are associated with environmental enteric dysfunction, and deficits in growth
27 and cognitive development.[1-4] Enteric infections and their sequelae disproportionately affect children
28 living in poverty in low- and middle-income countries (LMICs) due to inequitable access to healthcare,
29 inadequate water and sanitation infrastructure, and widespread fecal contamination of the environment.[2,
30 5-15]

31 Exposure to animal feces is an important transmission route of enteropathogens,[16-20] particularly
32 among children in LMICs where animals are ubiquitous and insufficient separation of animal feces from
33 domestic spaces is well documented.[20-26] Many pathogens capable of infecting humans are
34 transmissible via animal feces, some of which contribute significantly to the global burden of diarrheal
35 disease. Although the specific attributable fraction of animal-sourced infections is unquantified, four
36 pathogens that can be transmitted in animal feces (*Campylobacter* spp., *Cryptosporidium* spp.,
37 enteropathogenic *E. coli*, non-typhoidal *Salmonella*) are responsible for 28.3% of the estimated global
38 diarrhea deaths in children under five years.[16] Global animal feces production greatly exceeds that of
39 humans; livestock animal feces accounts for 80% of the global fecal load. A significant number of
40 households have these animals onsite, where human contact with feces is high.[27]

41 Significant challenges remain in understanding the scope of child exposure to zoonotic enteropathogens,
42 especially in the areas of highest risk. Such challenges are in part due to current approaches to exposure
43 assessment, which are diverse and distal from exposure itself.[28, 29] Researchers assess exposure
44 inconsistently with varied methods – including via survey, observation, and microbiology techniques –
45 which limits comparisons across studies and settings. Existing measures also overwhelmingly assess a
46 single attribute of exposure, typically related to animals (e.g., presence of animals) or the environment

47 (e.g., presence of animal feces).[28, 29] Such approaches fail to account for the multiple factors that lead
48 to exposure and do not capture more proximal factors (e.g., human contact with animals and animal feces)
49 that play a central role in exposure. For example, the presence of animals or animal feces may not
50 consistently be good proxies for exposure if children do not interact with the animals or their feces.
51 Conversely, accounting only for child behavior does not capture whether animal-sourced contaminants
52 are present. A standard measure that captures the multiple factors and conditions that contribute to
53 exposure is therefore needed to improve the assessment of child exposure to zoonotic enteropathogens
54 and to enable comparisons within and across communities.

55 In this study, we build upon existing measures of fecal-oral child exposure to zoonotic enteropathogens
56 and address prevailing measurement limitations by developing and validating a survey-based measure,
57 i.e., the FECEZ Enteropathogens Index. Here we report the sequential mixed methods approach we
58 employed to develop the index, followed by the results from our measurement development and
59 evaluation using data from northwestern coastal Ecuador.

60 **Methods**

61 **Defining and conceptualizing child exposure to zoonotic** 62 **enteropathogens**

63 Following established practices for measurement development,[30-32] we offer a preliminary definition
64 of our focal construct of interest, ‘child exposure to zoonotic enteropathogens’: ingestion of
65 enteropathogens through direct and indirect contact with animals, animal feces, and fecal contamination.
66 Alongside this definition, we developed a framework (Fig 1) adapted from the exposure science source-
67 to-outcome continuum,[33] as well as our recent qualitative research and systematic review, that is the
68 conceptual basis of the FECEZ Enteropathogens Index. The exposure science continuum – which
69 includes source; contaminant; contaminated environmental media, objects, and surfaces; behavior, route,

70 and outcome – delineates the specific elements that should be assessed to robustly measure exposure. Our
71 framework considers exposure to be constituted by two distinct sub-domains, which are the critical
72 attributes or characteristics of our focal construct. The *Environment* sub-domain includes the child’s
73 household, compound, and interpersonal environment. This sub-domain focuses on sources of zoonotic
74 enteropathogens (i.e., animals and their feces), the contaminant itself (i.e., enteropathogens in animal
75 feces), and contaminated environmental media, objects, and surfaces. The *Behavior* sub-domain includes
76 child behaviors inside and outside the household compound. This sub-domain focuses on interactions
77 with potential sources, contaminants, and contaminated environmental media, objects, and surfaces that
78 could lead to ingestion of zoonotic enteropathogens.

79 **Fig 1. Conceptual framework of the focal construct of interest, fecal-oral child exposure to zoonotic**
80 **enteropathogens.^a**

81 ^aConstruct: a well-defined and bounded subject of measurement; sub-domains: critical attributes or
82 characteristics of the construct of interest; dimensions: characteristics or elements that constitute the
83 construct of interest and its sub-domains [30, 31]

84 **Setting**

85 Our study was conducted in multiple communities in northwestern coastal Ecuador as part of an ongoing
86 birth cohort study, Enteropatógenos, Crecimiento, Microbioma, y Diarrea (referred to as ECoMiD).[34]
87 The EcoMiD study assesses the impact of environmental exposures on enteric pathogen infections, gut
88 microbiome composition, and development during the first two years of children’s lives. The region is
89 primarily populated by Afro-Ecuadorians, mestizos, and some indigenous individuals. Data were
90 collected in four ECoMiD sites: (1) Esmeraldas, the most urban site in the study area; (2) Borbón, a
91 smaller town in Esmeraldas Province that serves as a commercial center; (3) rural villages near Borbón
92 that are accessible by road; and (4) rural villages near Borbón that are accessible only by boat. Esmeraldas
93 (population: 160,000 [35]) is a densely populated city and the capital of Esmeraldas Province. The city

94 has the most access to water, sanitation, and infrastructure.[36] Borbón (population: 7,700 [35]), a town in
95 the Esmeraldas Province, is located at the convergence of the Cayapas, Santiago, and Onzole rivers. The
96 town has inadequate infrastructure for its size, including minimal water and sanitation infrastructure (e.g.,
97 untreated sewage, basic solid waste management).[36, 37] Approximately 125 small villages (population:
98 50-500 per village [35]) lie along the three rivers, some of which have access to Borbón via road (i.e.,
99 rural road communities) and have less developed infrastructure, though many have been recently
100 connected to drinking water systems.[36, 37] Other villages are only accessible by river (i.e., rural river
101 communities) and are comparatively more remote and lack centralized infrastructure.

102 **Overview of research design**

103 We used a sequential mixed methods approach [38] to create and evaluate the index, following measure
104 development best practices.[30-32] An overview of the three-phase approach is depicted in Fig 2 and
105 subsequently detailed.

106 **Fig 2. Schematic of the sequential mixed methods research design employed to create and evaluate** 107 **the FECEZ Enteropathogens Index to measure child exposure to zoonotic enteropathogens.^a**

108 ^aItems: survey questions that measure or reflect the construct of interest; content validity: the degree to
109 which the items (i.e., the measurement tool) adequately reflect the construct of interest; construct validity:
110 the degree to which the items accurately assess the construct of interest; test-retest reliability: the degree
111 to which the items consistently measure the construct of interest over time [30, 31]

112 **Phase 1: Item Development**

113 To identify content domains and generate potential index items (i.e., the survey questions that measure the
114 construct of interest), we conducted in-depth interviews (IDIs) and relied on our recently published
115 systematic review.[28] We finalized and evaluated the adequacy of the proposed items (i.e., content
116 validity) through expert review and cognitive interviews with individuals similar to our survey target
117 population, i.e., mothers of children 6-18 months of age in the study area.[30, 31]

118 **Phase 1 Stage 1a: Domain Identification**

119 In-depth go-along interviews

120 To define and conceptualize fecal-oral child exposure to zoonotic enteropathogens, we conducted 58 go-
121 along, semi-structured IDIs with mothers of children 6-18 months of age enrolled in the on-going birth
122 cohort study from January 21st to April 21st, 2021. Go-along IDIs combine interviewing and participant
123 observation, allowing the participant to actively engage with the spaces being discussed.[39, 40] We used
124 purposive quota sampling to interview equal proportions of mothers who did ($n=32$) and did not ($n=26$)
125 own animals to ensure a sufficient sample in each category to examine animal ownership's impact on
126 child exposure. Interviews queried conditions and maternal and child behaviors that could lead to child
127 exposure to zoonotic pathogens, capturing details related to animals, environmental conditions, and
128 behaviors on a typical day. To provide additional context about animals and environmental conditions, we
129 conducted 29 interviews with individuals from the same communities who were not part of the cohort
130 study and owned, cared for, and/or worked with various animals. Non-cohort IDIs queried how animals
131 are cared for (e.g., feeding, animal feces management) and by whom, and decision-making about animal
132 ownership and management. IDIs were conducted by co-author BCA, a qualitative researcher with more
133 than 10 years of research experience who grew up and lives in the study area.

134 We analyzed IDI data to understand the scope of child exposure to zoonotic enteropathogens, reviewing
135 transcript segments from relevant codes. We held debriefing meetings to conceptualize the constituents of
136 exposure and then recorded the prominence or frequency that each constituent occurred in the sample to
137 inform domain identification. Ultimately, we identified two content domains of exposure: the child's
138 environment and the child's behaviors inside the home and outside the home in the household compound.
139 Additional information about the qualitative methods, analyses, and findings are reported elsewhere.

140 Systematic review

141 To identify additional exposure attributes and characteristics and to make our index more generalizable to
142 other settings, we drew upon our recently published systematic review in which we audit existing

143 measures of human exposure to animal feces in LMICs. In our review, we identified 1,428 quantitative
144 measures and categorized them based on what type of information they captured (e.g., presence of
145 animals, animal fecal contamination of the environment, contact with animal feces) and where they fell
146 along the exposure science continuum. These categorizations were then used alongside findings from the
147 IDIs to solidify the content domains and dimensions and develop our conceptual framework. Detailed
148 information about the methods and results of the systematic review are described elsewhere.[28, 29]

149 **Phase 1 Stage 1b: Item Generation**

150 Using our conceptual framework based on the exposure science continuum, IDIs, and the systematic
151 review, we created an initial list of items for each sub-domain and dimension. We then reviewed existing
152 measures captured via the systematic review to identify additional relevant items, focusing on
153 comprehensiveness and alignment with our conception of child exposure.

154 **Phase 1 Stage 1c: Content Validity Assessment**

155 Expert review

156 To evaluate if proposed items adequately measure the construct (i.e., child exposure to zoonotic
157 enteropathogens), we conducted three rounds of expert review. First, four co-authors (BAC, MCF, RH,
158 KL) were sent draft items to assess the extent to which questions reflected the construct of interest and
159 were appropriately worded, ordered, representative, and comprehensive. Second, we conducted a
160 formalized expert review process with two other co-authors (JNSE, GOL) and two external experts who
161 were asked to evaluate each item on a scale of 1-4 for representativeness and clarity and provide
162 comments on representativeness, clarity, and comprehensiveness. Scores and comments for each item
163 were used to identify items to be edited, deleted, or added. Third, we translated items from English to
164 Spanish and the co-author (BCA) who conducted IDIs for this study reviewed and commented on items.

165 Cognitive interviews

166 As a final assessment of content validity, co-author BCA conducted 20 cognitive interviews with mothers
167 in the study area who were not enrolled in the ECoMiD study from March 2nd-21st, 2022. The researcher
168 asked participants to explain each question in their own words and how they determined their responses to
169 questions to evaluate the suitability for the target population and potential sources of response error. Edits
170 were made based on cognitive interview findings and debriefing meetings between co-authors AMB and
171 BCA.

172 **Phase 2: Index Development**

173 To collect data for index development and evaluation, we administered the revised items within a broader
174 survey, and then used principal components analysis (PCA) to reduce the number of items to maximize
175 parsimony.

176 **Phase 2 Stage 2a: Survey Administration**

177 Sampling frame

178 Although no consensus for optimal sample size exists for this cross sectional survey and subsequent
179 analysis, 200-300 participants has previously been found to be adequate for performing PCA.[30, 41-43]
180 Participant recruitment followed a random-walk method in the same neighborhoods where IDIs were
181 conducted and where the ECoMiD study participants reside. An enumerator walked through
182 neighborhoods and knocked on doors to screen participants for eligibility. To be eligible, the participant
183 was required to be aged 18 or older, a mother to a child six months to five years old, and not a member of
184 the ECoMiD cohort study. ECoMiD cohort participants were excluded to avoid research fatigue given
185 their ongoing participation in various cohort activities. If an eligible individual consented to participate,
186 the enumerator would administer the survey or make an appointment to return.

187 Data collection

188 The survey included the proposed measure items from Phase 1, as well as modules on mother and child
189 demographics, maternal perception and knowledge about exposure to animals and animal feces, water and

190 sanitation access, child health, and household characteristics. There were 52 proposed items designed to
191 measure the two exposure sub-domains. Each item asked mothers how often a particular event or
192 behavior occurred in the last week: never, rarely, sometimes, frequently (see S1 Table for full list of
193 items). Items were ordered to build in intensity or proximity to exposure. For example, *Environment*
194 items were asked first and items about child contact with animals and their feces were asked last. Items
195 that assessed the presence of animals and animal feces in the child's indoor and outdoor home
196 environment included potential follow-up items based on responses, meaning mothers were asked a range
197 of 52-105 questions (85 *Environment* items and 20 *Behavior* items). The survey was administered by a
198 single enumerator using Open Data Kit (ODK) and an Android tablet. Prior to data collection, the
199 enumerator completed one week of training by co-author AMB about the purpose of the survey, interview
200 techniques, research ethics, and logistics. Survey data collection occurred from August 3rd to September
201 30th, 2022. The total survey took an average of 30 minutes to complete. Participants received an
202 assortment of household items, such as soap and toothpaste, as compensation for their time.

203 To facilitate the assessment of test-retest reliability (i.e., how consistent index scores are across time), we
204 re-administered the survey sections with measure items to approximately 25% of participants within a
205 three to seven day window.[30, 31] Repeat surveys took approximately 15 minutes to complete and were
206 conducted among a random sample of those who agreed to a second visit. A 3-to-7-day window was
207 selected because the items refer to 'the past week' and we wanted to capture responses within an
208 overlapping period to reduce the likelihood of meaningful events or changes that would lead to different
209 responses to survey items.

210 **Phase 2 Stage 2b: Item Reduction**

211 Final sample assessment

212 Prior to PCA, we removed items if 1) few respondents (<5%) reported a certain condition or behavior
213 and 2) two items were highly correlated ($r > 0.90$). In the latter case, one item was retained and the other

214 was removed to reduce the number of initial items to a smaller subset of less-highly correlated items
215 where logic and theory supported items interrelatedness.

216 To evaluate if each set of items were suitable for PCA, we conducted three statistical tests. First, Kaiser-
217 Meyer-Olkin Measure of Sampling Adequacy [44] was used to determine the proportion of variance in
218 variables that may be caused by underlying factors. We considered values above 0.50 overall and per item
219 as adequate in demonstrating that PCA was useful to reduce the dimensionality of our data.[45, 46]
220 Second, Bartlett's Test of Sphericity [47] was used to test if the correlation matrix was an identity matrix,
221 meaning variables are unrelated and therefore not suitable for PCA. We considered p -values less than
222 0.05 an indication that the correlation matrix was not an identity matrix.[46] Lastly, we calculated the
223 determinant of the correlation matrix to assess if multicollinearity or singularity were issues in our data.
224 We considered determinant values >0.00001 to indicate no issues.[46]

225 Principal component analysis

226 To determine the optimal number of components (i.e., linear combinations of the original variables that
227 captures patterns of correlations related to an underlying factor and is derived through PCA [48]) that fit
228 each sub-domain, we conducted categorical PCA with multivariate analysis with optimal scaling using the
229 *princals* function from the *Gifi* package [49] in R Studio version 4.0.5 [50] on the raw ordinal data for the
230 two subdomains separately. We used a linearly scaled fit (i.e., linear knots with no interior knots) to
231 transform the ordinal values to be linearly scaled with equal distances between points, which aligns with
232 the meaning behind the ordinal data (i.e., the number of days). To identify the number of principal
233 components (PCs) to retain, we considered eigenvalues, a scree plot, and in part, theory based on our
234 conceptualization of exposure. We used Kaiser's Criteria (eigenvalues >1.0) and a scree plot to determine
235 the 'elbow' point, which demarks where the eigenvalues go from exponential decay to a linear trend.[31,
236 48, 51] To decide on final solutions, we iteratively re-ran analyses to examine solutions with a varying
237 number of components, balancing interpretability and the percent of variance explained to produce
238 parsimonious, functional, and interpretable indices.[30, 31, 52] We also assessed item loadings and

239 theoretical fit of each item within components. We decided *a priori* to conduct stepwise removal of items
240 with loadings <0.40 and/or that were loaded on several components and did not theoretically make
241 sense.[31, 48] After each item was removed, statistical tests and eigenvalues were re-assessed to make
242 sure the remaining data were still suitable for PCA and that we were still assessing an appropriate number
243 of components. The final component structures were assessed using knowledge of child exposure to
244 zoonotic enteropathogens to ensure that items and components were appropriate and relevant.

245 **Phase 3: Index Evaluation**

246 To create and evaluate the FECEZ Enteropathogens Index, we calculated scores based on PCA results and
247 assessed scores construct validity and test-retest reliability in R studio version 4.0.5.[50]

248 **Phase 3 Stage 3a: Index Item Scoring**

249 To calculate PC scores, we used an unweighted approach and calculated the sum of responses for each
250 final item using the original, ordinal values (i.e., 0 = never, 1 = rarely, 2 = sometimes, 3 = frequently).[30,
251 31] We calculated each components' score by summing each item's ordinal values together for items that
252 loaded to that component in the final PCA solutions. If items were cross-loaded, the item was considered
253 part of the component where the loadings were the largest.

254 To create index scores for substantive analysis, PC scores were used to calculate sub-domain (or sub-
255 index) and overall index scores. We used summations of ordinal data to calculate index scores, as
256 opposed to transformed scores from component loadings, to facilitate interpretability and index score
257 comparisons in future research given that loadings will differ by study population. Higher scores indicate
258 a greater frequency of occurrence.

259 **Phase 3 Stage 3b: Construct Validity and Test-retest Reliability Assessment**

260 To evaluate if the final index accurately assessed what it was designed to (i.e., construct validity), we
261 conducted bivariate linear regression analyses to examine if the measure behaves as expected in relation

262 to “known groups.”[30, 31] Specifically, we assessed whether scores for each component, sub-domain,
263 and the overall index were significantly different by community type (i.e., urban, commercial center, rural
264 road, rural river), hypothesizing that there would be detectable differences in exposure levels across the
265 urban-rural gradient because the number and diversity of animals varies across the sites. We also
266 investigated whether scores were significantly different by household animal ownership, as existing
267 literature suggest that exposure may be higher among children in households with animals, and by child
268 age, given that there are many developmental milestones that could impact exposure in the first five years
269 of life.

270 To determine if the index provided a stable measure that can be used on repeat occasions (i.e., test-retest
271 reliability),[30, 31] we calculated intraclass correlation coefficient (ICC) estimates and their 95%
272 confidence intervals using the *ICC* function from the *psych* package [53] in R studio version 4.0.5[50]
273 based on a single-rating, absolute agreement, two-way mixed effects model.[54] ICCs were calculated for
274 each PC, the *Environment* and *Behavior* sub-indices, and the overall index for participants who were
275 surveyed twice. We used the following guidelines to evaluate ICC values: <0.50 poor reliability, 0.50-
276 0.75 moderate reliability, 0.76-0.90 good reliability, and >0.90 excellent reliability.[54]

277 **Ethics**

278 Emory University (IRB # 00101202) and Universidad San Francisco de Quito Institutional (IRB # 2018-
279 022M and 021-011M) Review Boards approved all study activities. Participants provided written consent
280 prior to data collection.

281 **Results**

282 **Participant Demographics**

283 In total, we administered 297 surveys across the four study sites. Most households owned at least one type
284 of animal (55.6%, $n=165$), with dogs and cats being the most common (Table 1). A quarter ($n=75$) of

285 children were reported by mothers to have had a fever in the week prior, 13.5% ($n=40$) had diarrhea, and
 286 7.7% ($n=23$) had vomited. Sex-disaggregated demographic characteristics are provided in S2 Table.

287 **Table 1. Maternal, child, and household characteristics for total sample and by the four study sites**
 288 **($n=297$)**

Characteristics	Total		Rural river		River road		Commercial center		Urban	
	<i>n</i>		<i>n</i>		<i>n</i>		<i>n</i>		<i>n</i>	
Number of participants	297		46	15.4%	76	25.6%	98	33.0%	77	25.9%
Maternal characteristics										
Age (mean [std] in years)	29	(8.0)	28	(7.0)	30	(9.0)	28	(7.0)	32	(8.0)
<i>Ethnicity</i>										
Afro-Ecuadorian	221	74.4%	42	91.3%	64	84.2%	67	68.4%	48	62.4%
Mestizo	70	23.6%	4	8.7%	12	15.8%	26	26.5%	28	36.4%
Indigenous - Chachi	2	0.7%	0	0.0%	0	0.0%	2	2.0%	0	0.0%
Other	4	1.3%	0	0.0%	0	0.0%	3	3.1%	1	1.3%
Education (mean [std] in years)	11.5	(3.5)	9	(4.0)	11	(3.5)	12	(3.0)	13	(3.0)
Child characteristics										
Age (mean [std] in months)	33	(15.5)	34	(16.0)	36	(15.0)	34	(15.0)	29	(16.0)
Sex- female	153	51.5%	25	54.3%	44	57.9%	48	49.0%	36	46.8%
Currently breastfed	34	11.4%	3	6.5%	4	5.3%	9	9.2%	18	23.4%
<i>Symptoms in last 7 days</i>										
Diarrhea	40	13.5%	11	23.9%	11	14.5%	14	14.3%	4	5.2%
Fever	75	25.3%	19	41.3%	30	39.5%	19	19.4%	7	9.1%
Vomit	23	7.7%	3	6.5%	8	10.5%	10	10.2%	2	2.6%
Blood in stool	1	0.3%	0	0.0%	0	0.0%	1	1.0%	0	0.0%
Household characteristics										
Number of people* (mean [std])	5	(2.5)	6	(2.0)	5	(2.0)	5	(3.0)	5	(3.0)
<i>Owens animal(s)</i>										
Dogs	165	55.6%	22	47.8%	48	63.2%	61	62.2%	34	44.2%
Cats	115	38.7%	13	28.3%	29	38.2%	44	44.9%	29	37.7%
Cats	62	20.9%	8	17.4%	20	26.3%	22	22.4%	12	15.6%
Free-range chickens	35	11.8%	4	8.7%	14	18.4%	15	15.3%	2	2.6%
Ducks	3	1.0%	0	0.0%	3	3.9%	0	0.0%	0	0.0%
Dairy cattle	2	0.7%	0	0.0%	0	0.0%	2	2.0%	0	0.0%
Horses	1	0.3%	0	0.0%	1	1.3%	0	0.0%	0	0.0%
Pigs	12	4.0%	2	4.3%	5	6.6%	5	5.1%	0	0.0%
Rabbits	7	2.4%	0	0.0%	5	6.6%	2	2.0%	0	0.0%
<i>Animal(s) present</i>	290	97.6%	46	100.0%	75	98.7%	98	100.0%	71	92.2%
<i>Animal feces present</i>	210	70.7%	38	82.6%	44	57.9%	79	80.6%	49	63.6%

289 * $n=292$, Five observations have missing values

290

291 Index Development

292 Final Sample Assessment

293 The survey included 105 items for potential inclusion in the final measures: 85 *Environment* items and 20
 294 *Behavior* items. All participants responded “never” to 37 *Environment* items and one *Behavior* item, so
 295 they were eliminated due to their irrelevancy to this population. We eliminated 21 additional *Environment*
 296 items because they were near zero variance predictors, indicating they had limited relevance for the

297 sample. Lastly, two *Behavior* items were eliminated because they were highly correlated and theoretically
298 similar to two retained items. Specifically, items that captured children playing in sand outside the house
299 and children putting sand in their mouths were removed because they were highly correlated with
300 children playing in dirt or soil outside the house and children putting dirt or soil in their mouths ($r>0.90$).
301 PCA was therefore conducted with 27 *Environment* items and 17 *Behavior* items to create two sub-
302 indices. Distributions of item responses are in S3 Table. Items that were omitted and reasons for omission
303 are in S1 Table.

304 **Principal Component Analysis**

305 Our three pre-analysis tests indicated that remaining data for each sub-domain were suitable for PCA. For
306 the *Environment* and *Behavior* items, the Kaiser-Meyer-Olkin measure values were 0.71 and 0.75,
307 respectively, indicating acceptable sampling adequacy. The Bartlett's Test of Sphericity revealed that the
308 between-item correlations were sufficient for PCA for both sets of items (*Environment* items: K-squared
309 = 2094.8, degrees of freedom (df) = 19, p -value <0.001; *Behavior* items: K-square = 2079.4, df = 13, p -
310 value <0.001). There were also no issues with multicollinearity or singularity; determinants of the
311 correlation matrices were 0.002 and 0.01 for the *Environment* and *Behavior* items, respectively.

312 **Child Environment sub-domain**

313 For the *Environment* sub-domain, we determined that a five-component solution best suited the data
314 theoretically, based on a screeplot, Kaiser's Rule (eigenvalues >1.0), and the amount of variance
315 explained by PCs. Seven *Environment* items were omitted during analyses due to loadings <0.40 and/or
316 cross-loading that was not interpretable (S1 Table). The final *Environment* sub-index included 20 items
317 and explained 57% of the variance (Table 2). The fit appeared to be good based on the interpretability of
318 the PCs, a loss value (i.e., a measure of goodness of fit) of 0.89, and a solution being obtained with 43
319 iterations. Of note, lower loss values and fewer iterations indicate better fits, though no exact thresholds
320 exist.

321 **Table 2. Eigen values, explained variance, cumulative explained variance, and component loadings**
 322 **for child *Environment* sub-domain PCA solution**

Solution characteristics	PC1	PC2	PC3	PC4	PC5
Eigen value	3.76	2.35	2.21	1.73	1.26
Variance explained by PC	0.19	0.12	0.11	0.09	0.06
Cumulative variance explained	0.19	0.31	0.42	0.51	0.57
Solution items					
Mother personally feeds or gives water to an animal	-0.79	-0.08	-0.13	-0.02	-0.03
Mother personally cleans the habitat or place where an animal sleeps and/or defecates	-0.75	0.09	-0.26	0.02	-0.19
Mother personally bathes, cleans, or grooms an animal	-0.71	0.12	-0.33	-0.00	-0.05
Mother personally touches or play with an animal	-0.70	0.07	-0.30	0.00	0.03
Mother personally eliminates or cleans the poop of an animal	-0.70	-0.12	-0.29	0.15	-0.24
Dogs enter the house	-0.53	0.30	-0.03	-0.28	0.32
Free-range chickens spend time outside near the house	-0.27	-0.77	0.12	-0.06	0.00
Free-range chickens enter the house	-0.12	-0.72	0.05	-0.05	-0.01
Free-range chicken poop outside the house near or in the yard	-0.24	-0.70	0.20	0.11	0.03
Free-range chicken poop inside the house	-0.10	-0.56	0.11	0.15	0.02
Cats enter the house	-0.37	0.22	0.76	0.23	0.02
Cats spend time outside near the house	-0.25	0.22	0.64	0.29	-0.13
Cats sleep inside the house	-0.40	0.21	0.60	0.16	0.11
Cat poop outside the house near or in the yard	-0.15	-0.05	0.40	0.30	-0.26
Dairy cattle spend time outside near the house	-0.10	-0.11	0.26	-0.70	-0.05
Dairy cattle poop outside the house near or in the yard	-0.00	-0.02	0.25	-0.70	-0.08
Household member apart from mother and child under 5 years works or cares for an animal	-0.16	-0.08	0.37	-0.52	0.00
Dogs sleep inside the house	-0.44	0.23	0.06	-0.12	0.58
Dog poop outside the house near or in the yard	-0.17	0.24	0.07	-0.22	-0.57
Dogs spend time outside near the house	0.04	0.17	-0.01	-0.13	-0.53

323 PC = Principal component; bold numeric values indicate item loading to the specific PC

324
 325 The *Environment* sub-domain PCA yielded strong loadings onto five interpretable components, each
 326 listed with the proportion of variance accounted for: maternal factors (19%), free-range chicken factors
 327 (12%), cat factors (11%), dairy cattle factors (9%), and dog factors (6%) (Table 2). The PCs broadly
 328 corresponded to our two initially hypothesized *Environment*-related dimensions: the child’s household
 329 and compound environment and the child’s interpersonal environment. Specifically, PCs 2-5 included
 330 items related to specific species of animals and their feces. The item about other household members
 331 working with or caring for animals loaded on PC4 with dairy cattle items, which could represent farming
 332 communities/households where dairy cattle are present and family’s own, work with, and/or care for

333 them. PC1 included items about maternal behaviors and interactions with animals and their feces. The
 334 items about dogs entering and dogs and cats sleeping in the household also loaded on this component,
 335 which likely indicates that mothers interact with dogs and cats and their feces that enter or sleep inside
 336 their house. Dogs entering the household loaded to PC1 (maternal factors) and not to PC5 (dog factors),
 337 which could be indicative of dogs specifically entering houses to be fed, bathed, groomed, or played with
 338 and contributing to interpersonal environmental contamination.

339 **Child Behavior sub-domain**

340 For the *Behavior* sub-domain, a two-component solution best suited the data. Three additional items were
 341 omitted during analyses due to loadings <0.40 and/or cross-loading (S1 Table). The final *Behavior* sub-
 342 index included 14 items, explaining 42% of the variance (Table 3). The fit appeared adequate with a loss
 343 value of 0.79 and a solution obtained with 12 iterations, and based on the PCs interpretability.

344 **Table 3. Eigen values, explained variance, cumulative explained variance, and component loadings**
 345 **for child *Behavior* sub-domain PCA solution**

Solution characteristics	PC1	PC2
Eigen value	3.34	2.52
Variance explained by PC	0.24	0.18
Cumulative variance explained	0.24	0.42
Solution items		
Child puts objects or toys that had contact with the dirt outside your house in their mouth	0.73	-0.25
Child plays outside the house without shoes on	0.67	0.13
Child puts dirt or soil in their mouth	0.59	-0.28
Child plays in dirt or soil outside the house	0.59	0.05
Child puts objects or toys that had contact with the floor inside your house in their mouth	0.58	-0.20
Child plays outside the house in an area where an animal lives or sleeps	0.56	0.12
Child puts shoes in their mouth	0.53	-0.32
Child plays with or carries around shoes like a toy	0.52	0.02
Child touches or plays with an animal	0.45	0.34
Child cleans or helps others clean the habitat or place where an animal sleeps and/or defecates	0.16	0.76
Child bathes, cleans, or grooms or helps others bathe, clean, or groom an animal	0.22	0.75
Child feeds or gives water or helps others feed or give water to an animal	0.29	0.67
Child cares for or helps others care for an animal that was sick	0.02	0.56
Child touches, removes, or cleans animal poop	0.11	0.54

346 PC = Principal component; bold numeric values indicate item loading to the specific PC

347 The *Behavior* sub-domain PCA yielded strong loadings onto two interpretable components, each listed
348 with the proportion of variance accounted for: play and mouthing behaviors (24%), and animal caregiving
349 and feces management behaviors (18%) (Table 3). The components broadly corresponded to our
350 structuring of questions that build in proximity to exposure to animals and their feces. Specifically, PC1 is
351 comparatively more distal from exposure, including items about child play in potentially risky
352 environments and mouthing potentially contaminated objects. PC2 includes items of increasing proximity
353 to exposure, specifically caring for or helping others care for animals and interacting with animal feces.

354 **Index Evaluation**

355 **Index Item Scoring**

356 The mean overall FECEZ Enteropathogens Index score was 27.21 (standard deviation [SD]: 12.75) out of
357 102 (Table 4). The average *Environment* sub-domain and *Behavior* sub-domain scores were 14.63 (SD:
358 8.57) out of 60 and 12.57 (SD: 7.10) out of 42, respectively. *Environment* and *Behavior* sub-domain
359 scores were moderately positively associated, $r(295) = 0.31$, $p < 0.01$ (S2 Fig). Histograms for sub-domain
360 and overall index scores are provided in S1-S3 Figs. On average, scores were highest among children
361 living in the rural river study site (Table 5). Children in the urban study site had the lowest average scores
362 except for *Environment* PC5 (dog factors), signifying less child interaction with animals and their
363 environment and less presence of animals and their feces, apart from dogs.

364 **Table 4. Means and standard deviations of scores for Principal Components (PCs), Child**
365 ***Environment* and *Behavior* sub-domains, and the overall FECEZ Enteropathogens Index (n=297)**

	No. of items	Possible Score range	Mean score (std)
Total	34	0-102	27.21 (12.75)
<i>Child Environment sub-total</i>	20	0-60	14.63 (8.57)
PC1: Maternal factors	6	0-18	5.42 (5.64)
PC2: Free-range chicken factors	4	0-12	1.48 (2.67)
PC3: Cat factors	4	0-12	2.92 (3.35)
PC4: Dairy cattle factors	3	0-9	0.30 (0.96)
PC5: Dog factors	3	0-9	4.52 (1.98)
<i>Child Behavior sub-total</i>	14	0-42	12.57 (7.10)
PC1: Play and mouthing behaviors	9	0-27	11.99 (6.55)
PC2: Animal caregiving and feces behaviors	5	0-15	0.58 (1.72)

366

367 **Validity and Reliability Assessment**

368 The PC, sub-domain, and total index scores differed significantly by community type (Table 5),
369 indicating good construct validity (i.e., known-groups validity). As hypothesized, there was a statistically
370 significant difference between scores across study sites. Children in rural river and rural road
371 communities and the commercial center had significantly higher *Environment* sub-domain, *Behavior* sub-
372 domain, and total index exposure scores compared to urban-residing children. There were also significant
373 differences across PCs, though these differences varied by component as expected. For example, children
374 in rural river and rural road communities and the commercial center had significantly higher scores
375 related to free-range chickens (i.e., *Environment* PC2) compared to urban-residing children, which was
376 expected given the sparseness of free-range chickens in the urban study site. However, only children in
377 the commercial center study site had significantly higher scores related to dogs (i.e., *Environment* PC5),
378 indicative of the presence of dogs and their feces across all study sites.

Table 5. Component, sub-domain, and overall FECEZ Enteropathogens Index scores by community type and animal ownership, and child age

	Mean score (sd) by community type				Mean score (sd) by animal ownership			Mean score (sd) by child age in months (mths)			
	Urban (n=77, ref.)	Commercial center (n=98)	Rural road (n=76)	Rural river (n=46)	None (n=132, ref.)	1 type (n=110)	>1 type (n=55)	6-12 mths (n=39, ref.)	13-18 mths (n=22)	19-24 mths (n=43)	25-60 mths (n=193)
Total	18.10 (11.55)	29.07 (11.45)	29.82 (11.12)	34.15 (12.05)	20.52 (11.16)	30.05 (10.63)	37.55 (11.25)	19.64 (11.98)	30.50 (12.63)	27.79 (13.02)	28.23 (12.40)
<i>Child Environment sub-domain</i>	10.32 (8.19)	15.82 (8.66)	16.28 (8.13)	16.61 (7.40)	9.04 (6.06)	16.91 (6.69)	23.51 (7.50)	11.87 (8.44)	14.82 (7.90)	15.35 (9.22)	15.01 (8.48)
PC1: Maternal factors	4.14 (5.85)	4.98 (5.02)	6.86 (5.67)	6.15 (6.03)	2.02 (3.23)	7.54 (5.52)	9.36 (5.87)	3.97 (5.35)	5.64 (6.12)	5.12 (5.56)	5.76 (5.64)
PC2: Free-range chicken factors	0.16 (0.96)	1.39 (2.32)	2.03 (3.27)	2.96 (3.19)	0.85 (2.24)	1.63 (2.63)	2.67 (3.23)	1.18 (2.55)	1.95 (2.84)	1.49 (2.59)	1.48 (2.71)
PC3: Cat factors	1.64 (2.79)	3.54 (3.22)	3.43 (3.82)	2.87 (3.18)	2.04 (2.56)	2.62 (3.35)	5.62 (3.70)	1.95 (2.86)	2.05 (2.50)	4.02 (3.61)	2.96 (3.41)
PC4: Dairy cattle factors	0.00 (0.00)	0.89 (1.49)	0.04 (0.34)	0.00 (0.00)	0.09 (0.47)	0.25 (0.84)	0.91 (1.62)	0.05 (0.32)	0.55 (1.14)	0.23 (0.72)	0.34 (1.06)
PC5: Dog factors	4.39 (2.23)	5.02 (1.51)	3.92 (1.94)	4.63 (2.25)	4.04 (1.97)	4.87 (1.92)	4.95 (1.91)	4.72 (2.18)	4.64 (1.89)	4.49 (2.15)	4.47 (1.92)
<i>Child Behavior sub-domain</i>	7.78 (5.19)	13.26 (6.07)	13.54 (7.19)	17.52 (7.28)	11.48 (7.10)	13.15 (6.89)	14.04 (7.22)	7.77 (7.58)	15.68 (7.82)	12.44 (6.56)	13.22 (6.63)
PC1: Play and mouthing behaviors	7.65 (4.92)	12.36 (5.55)	12.92 (6.42)	16.96 (6.88)	11.35 (6.96)	12.40 (6.22)	12.73 (6.17)	7.67 (7.46)	15.18 (6.89)	12.23 (6.41)	12.45 (5.99)
PC2: Animal caregiving and feces behaviors	0.13 (0.77)	0.90 (2.01)	0.62 (1.84)	0.59 (1.86)	0.14 (0.71)	0.75 (1.68)	1.31 (2.85)	0.100 (0.50)	0.50 (1.92)	0.21 (0.71)	0.77 (1.97)

Bold numeric values indicate p -value ≤ 0.05 for bivariate linear regression between scores and community type, animal ownership, and child age

379 Children residing in households that owned at least one type of animal had significantly higher
380 *Environment* sub-domain, *Behavior* sub-domain, and total index exposure scores compared to children in
381 households that did not own any animals, as hypothesized. There were significant differences across PCs,
382 though differences varied by component, as expected. For example, children in households that owned
383 animals had significantly higher scores related to maternal factors (i.e., *Environment* PC1) compared to
384 those residing in households without animals, indicative of the absence of animals in the household for
385 mothers to interact with and care for. Conversely, there was not a significant difference between child
386 play and mouthing scores (i.e., *Behavior* PC1), which was expected given that child play in potentially
387 risky environments and mouthing potentially contaminated objects is not dependent on household animal
388 ownership.

389 There were minimal differences in *Environment* sub-domain scores by child age, and there were
390 significant differences between the *Behavior* sub-domain and *Behavior*-related PCs, as expected. For
391 example, children who were 25-60 months of age had significantly higher scores related to animal
392 caregiving and feces behavior (i.e., *Behavior* PC2) compared to children 6-12 months of age, which is
393 expected given child developmental phases. There were also significant differences in overall exposure
394 scores by child age, as hypothesized. Children become more mobile and active as they age, increasing
395 events that could contribute to exposure.

396 ICC values, based on data from participants who were surveyed twice ($n=66$), for sub-domain and the
397 overall index indicated good test-retest reliability (Table 6). The *Environment* sub-index, *Behavior* sub-
398 index, and overall index had ICC values of 0.79, 0.81, and 0.81, respectively. ICC values for PCs showed
399 moderate to excellent reliability. All five *Environment* sub-domain PCs had moderate to excellent test-
400 retest reliability. For the *Behavior* sub-index, PC1 had good reliability (ICC: 0.86) and PC2 had moderate
401 reliability (ICC: 0.66).

402 **Table 6. Component, sub-domain, and overall FECEZ Enteropathogens Index scores intraclass**
403 **correlation coefficient [ICC] estimates**

	ICC estimates
	Kappa (95% CI)
Total	0.81 (0.70, 0.88)
<i>Child Environment sub-domain</i>	0.79 (0.68, 0.86)
PC1: Maternal factors	0.87 (0.80, 0.92)
PC2: Free-range chicken factors	0.75 (0.63, 0.84)
PC3: Cat factors	0.83 (0.73, 0.89)
PC4: Dairy cattle factors	0.51 (0.31, 0.67)
PC5: Dog factors	0.70 (0.54, 0.81)
<i>Child Behavior sub-domain</i>	0.81 (0.70, 0.88)
PC1: Play and mouthing behaviors	0.86 (0.79, 0.92)
PC2: Animal caregiving and feces behaviors	0.66 (0.49, 0.78)

404

405 **Discussion**

406 The FECEZ Enteropathogens Index – a valid and reliable two-domain, 34-item survey-based measure –
407 advances our ability to assess exposure by quantifying the multiple components of child exposure to
408 zoonotic enteropathogens with higher resolution. The measure produces a composite, continuous value
409 that allows for the assessment of relative intensity of and variation in exposure and may decrease the
410 likelihood of exposure misclassification. This tool is an initial step towards developing a universal index
411 to compare exposure across populations. Below we describe the major strengths of the FECEZ
412 Enteropathogens Index, including its integration of multiple factors to assess exposure and its ability to
413 quantify the degree of exposure. We then outline the implications for research and practice, offering
414 recommendations for the application of our tool and for future research to further refine the measure to
415 ensure applicability to a broader context.

416 The FECEZ Enteropathogens Index revealed that exposure to zoonotic enteropathogens was ubiquitous
417 among children in northwestern coastal Ecuador, a finding that would have been masked with the use of
418 commonly used measures. Specifically, only two children were not exposed (i.e., had a score of zero). In
419 contrast, if we had used animal ownership or the presence of animal feces as a measure of exposure in this
420 study, 44% ($n=132$) and 33% ($n=87$) of children would have been classified as having no exposure,

421 respectively. The assessment of child exposure using the presence of animals would have been more
422 closely aligned the FECEZ index measure, but merely noting the presence of animals provides
423 substantially less information compared to our index, which captures many factors to comprehensively
424 assess exposure. Importantly, our index includes an entire domain with items to assess child behavior – a
425 novel feature given that only 9% of existing exposure measures identified by our systematic review
426 incorporated human behavior [28] – as well as multiple environmental factors, which are potential
427 sources of enteropathogens (e.g., presence of specific animals and their feces) and pre-requisites of
428 exposure.

429 The FECEZ Enteropathogens Index also captured exposure heterogeneity among our study population, a
430 level of detail that would have been missed with typical binary measures of exposure. We found
431 significant differences in the degree of exposure across communities and by household animal ownership
432 and age. Commonly used binary measurements of zoonotic exposure [28] do not capture this variation
433 and provide less information than a multidimensional measure like our FECEZ Enteropathogens Index.
434 For example, binary measures such as the presence of animal feces or household animal ownership are
435 unable to identify risk gradients among children and have the potential to introduce bias by misclassifying
436 some children as unexposed when they in fact have some level of exposure.[55]

437 **Implications for research and practices**

438 Child exposure scores produced using the FECEZ Enteropathogens Index were reliable and valid,
439 suggesting that the measure functions well and can be tested more broadly. We found that the domain-
440 specific and overall indices had good test-retest reliability, and exposure scores were significantly
441 different across known-groups, demonstrating construct validity. Expert review and cognitive interviews
442 during the item development phase strengthened the index’s ability to adequately measure exposure (i.e.,
443 content validity). Engaging urban and rural communities with different types of animals and husbandry
444 practices increases the final measure’s generalizability, as does our use of existing exposure measures to
445 create and evaluate the index.

446 Researchers and practitioners across sectors can use the 34-item index that we developed in this study to
447 assess child exposure in several ways. For example, the index can inform program design by assessing
448 baseline levels of exposure by sub-domain to identify areas for targeted intervention. The sub-domains
449 can also be combined to comprehensively measure child exposure and identify inequities and vulnerable
450 sub-populations within communities. Throughout program implementation, measurement using this index
451 can enable the monitoring of exposure trends over time with higher resolution, supplying critical data for
452 the development and evaluation of interventions. Such data can, in turn, facilitate the identification of
453 upstream determinants of exposure and the assessment of the consequences of exposure among children.

454 Prior to implementing the FECEZ Enteropathogens Index in new locations, we offer two specific
455 recommendations for researchers and practitioners. These steps are necessary for the measure's use until
456 its reliability and validity have been demonstrated across diverse contexts. However, researchers and
457 practitioners may use the current version of the measure if the context is similar to northwestern coastal
458 Ecuador, or use the index as a starting point for other local contexts.[56] First, we suggest evaluating
459 content validity through cognitive interviews. During interviews, participants should be queried about the
460 presence of feces from the types of animals that are relevant to the research context and known to be of
461 high concern for pathogens transmitted in animal feces.[16] This step is crucial for ensuring a
462 comprehensive assessment of exposure across settings because the FECEZ Enteropathogens Index
463 consists of causal or formative indicators (as indices do generally [31, 57, 58]), meaning the absence of
464 specific types of animals in this Ecuadorian study does not mean that they are not part of the construct,
465 'child exposure to zoonotic enteropathogens.' For example, we found that dogs, dairy cattle, cats, and free-
466 range chickens were contributing to child exposure in Ecuador, which are all known to transmit four of
467 the five pathogens that have been identified as of highest concern for pathogens transmitted in animal
468 feces (i.e., *Campylobacter* spp., non-typhoidal *Salmonella*, *Cryptosporidium*, and *Toxoplasma*
469 *gondii*).[16] However, other animals can also transmit these pathogens (e.g., swine can transmit non-
470 typhoidal *Salmonella* and *Cryptosporidium*), and many other pathogens may have significant

471 transmission potential through various types of animal feces (e.g., shiga toxin *E. coli*, *Toxocara*
472 *canis/Toxocara cati*).[16] A list of animals to consider based on our research and existing literature is
473 publicly available on OSF.[59]

474 Second, when applying the index in new locations, we recommend that researchers and practitioners
475 conduct a PCA to test the component structure and assess test-retest reliability and content and construct
476 validity. Such testing will explore the index’s external validity and help with progress toward the
477 development of a global index. Survey items, scoring instructions, and detailed recommendations for the
478 assessment of reliability and validity are publicly available on OSF.[59]

479 **Strengths and Limitations**

480 A strength of this research is its’ multi-phase, rigorous approach to measurement development and
481 evaluation. Use of qualitative interviews, a systematic review, expert review, and cognitive interviews for
482 item development strengthen the measure’s content validity. Items serve as proxies to capture the
483 components of exposure and many exposure pathways, improving upon existing exposure measures that
484 only assess one aspect of exposure. However, this measure is limited by its inability to be a proxy for
485 every potential animal feces exposure pathway. The measure does not directly assess pathways related to
486 food, flies, and fluids (water), which are challenging to assess through survey and were therefore not
487 included. Still, assessment of the final measure demonstrates good construct validity and test-retest
488 reliability. Additionally, we were able to collect data from four study sites that differ, for example, by
489 size, infrastructure, and livelihood systems. The diverse sample strengthens the external validity of the
490 final measure, although testing in different populations in varied geographic locations is needed to further
491 assess the measure’s generalizability. A limitation is our inability to examine how the measure performs
492 compared to existing measures (i.e., concurrent validity), which we were unable to do because there is not
493 a “gold standard” measure of child exposure to zoonotic enteropathogens. Lastly, our unweighted scoring
494 approach may be less precise than a weighted approach that allows item scores to be based on their
495 contribution to the component or factor. However, studies have found that weighting only improves

496 precision moderately, makes interpretation challenging, and inhibits across-study comparisons.[31, 48]
497 Producing unweighted scores is more user friendly, and will facilitate the standardization of exposure
498 measurement and comparison of scores across studies.

499 **Conclusion**

500 The FECEZ Enteropathogens Index advances our ability to assess exposure by quantifying the multiple
501 components of child exposure to zoonotic enteropathogens with higher resolution. Researchers and
502 practitioners may use the current version of the measure if the context is similar to the current study.
503 Additional rigorous testing and evaluation of the index in different settings would ensure its reliability,
504 validity, and cross-cultural equivalence in other contexts. Cognitive interviews and field testing of the tool
505 should be conducted in diverse geographic locations as part of the ongoing iterative process aimed at
506 establishing a global, standard exposure measure. The FECEZ Enteropathogens Index will be suitable for
507 widespread use once its performance has been evaluated across multiple settings.[56] The advent of such
508 a measure could impact on our understanding of child zoonotic enteropathogen exposures and can offer
509 valuable evidence for the development of community's public health and policy agendas.

510 **Acknowledgements**

511 We would like to thank local field staff, especially in Borbón, for their help with recruiting, logistics, and
512 data collection for this project; and Nick Laramée and Jayden Pace Gallagher for their contributions to the
513 qualitative and systematic review phases of this study; and all the participants in Ecuador that made this
514 project possible. We also thank Drs. Frederica Lamar and Sarah McKune for serving as external
515 reviewers for the formalized expert review process.

516

517

518

519

520 **References**

- 521 1. Barredo L, Agyepong I, Liu G, Reddy S. Ensure healthy lives and promote well-being for all at
522 all ages. UN Chronicle. 2015;51(4):9-10.
- 523 2. Hug L, Alexander M, You D, Alkema L, for Child UI-aG. National, regional, and global levels
524 and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a
525 systematic analysis. The Lancet Global Health. 2019;7(6):e710-e20.
- 526 3. Lozano R, Fullman N, Abate D, Abay SM, Abbafati C, Abbasi N, et al. Measuring progress from
527 1990 to 2017 and projecting attainment to 2030 of the health-related Sustainable Development Goals for
528 195 countries and territories: a systematic analysis for the Global Burden of Disease Study 2017. The
529 lancet. 2018;392(10159):2091-138.
- 530 4. Paulson KR, Kamath AM, Alam T, Bienhoff K, Abady GG, Abbas J, et al. Global, regional, and
531 national progress towards Sustainable Development Goal 3.2 for neonatal and child health: all-cause and
532 cause-specific mortality findings from the Global Burden of Disease Study 2019. The Lancet.
533 2021;398(10303):870-905.
- 534 5. Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and
535 national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a
536 systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392(10159):1736-88.
- 537 6. Reiner RC, Wiens KE, Deshpande A, Baumann MM, Lindstedt PA, Blacker BF, et al. Mapping
538 geographical inequalities in childhood diarrhoeal morbidity and mortality in low-income and middle-
539 income countries, 2000–17: analysis for the Global Burden of Disease Study 2017. The Lancet.
540 2020;395(10239):1779-801.
- 541 7. Prüss-Ustün A, Wolf J, Bartram J, Clasen T, Cumming O, Freeman MC, et al. Burden of disease
542 from inadequate water, sanitation and hygiene for selected adverse health outcomes: an updated analysis
543 with a focus on low-and middle-income countries. International journal of hygiene and environmental
544 health. 2019;222(5):765-77.

- 545 8. Walker CLF, Rudan I, Liu L, Nair H, Theodoratou E, Bhutta ZA, et al. Global burden of
546 childhood pneumonia and diarrhoea. *Lancet* (London, England). 2013;381(9875):1405-16.
- 547 9. Perin J, Mulick A, Yeung D, Villavicencio F, Lopez G, Strong KL, et al. Global, regional, and
548 national causes of under-5 mortality in 2000–19: an updated systematic analysis with implications for the
549 Sustainable Development Goals. *The Lancet Child & Adolescent Health*. 2022;6(2):106-15.
- 550 10. Lunn P. The impact of infection and nutrition on gut function and growth in childhood. *Proc Nutr*
551 *Soc*. 2000;59:147.
- 552 11. Lunn PG. Growth retardation and stunting of children in developing countries. *British Journal of*
553 *Nutrition*. 2002;88(2):109-10.
- 554 12. Guerrant RL, Oria RB, Moore SR, Oria MO, Lima AA. Malnutrition as an enteric infectious
555 disease with long-term effects on child development. *Nutr Rev*. 2008;66(9):487-505.
- 556 13. Guerrant RL, DeBoer MD, Moore SR, Scharf RJ, Lima AA. The impoverished gut--a triple
557 burden of diarrhoea, stunting and chronic disease. *Nature reviews Gastroenterology & hepatology*.
558 2013;10(4):220-9.
- 559 14. Keusch GT, Denno DM, Black RE, Duggan C, Guerrant RL, Lavery JV, et al. Environmental
560 Enteric Dysfunction: Pathogenesis, Diagnosis, and Clinical Consequences. *Clin Infect Dis*. 592014. p.
561 S207-12.
- 562 15. Petri WA, Jr., Miller M, Binder HJ, Levine MM, Dillingham R, Guerrant RL. Enteric infections,
563 diarrhea, and their impact on function and development. *The Journal of clinical investigation*.
564 2008;118(4):1277-90.
- 565 16. Delahoy MJ, Wodnik B, McAliley L, Penakalapati G, Swarhout J, Freeman MC, et al. Pathogens
566 transmitted in animal feces in low- and middle-income countries. *International Journal of Hygiene and*
567 *Environmental Health*. 2018;221(4):661-76.
- 568 17. Penakalapati G, Swarhout J, Delahoy MJ, McAliley L, Wodnik B, Levy K, et al. Exposure to
569 Animal Feces and Human Health: A Systematic Review and Proposed Research Priorities. *Environ Sci*
570 *Technol*. 2017;51(20):11537-52.

- 571 18. Prendergast AJ, Gharpure R, Mor S, Viney M, Dube K, Lello J, et al. Putting the “A” into WaSH:
572 a call for integrated management of water, animals, sanitation, and hygiene. *The Lancet Planetary Health*.
573 2019;3(8):e336-e7.
- 574 19. Pickering AJ, Null C, Winch PJ, Mangwadu G, Arnold BF, Prendergast AJ, et al. The WASH
575 Benefits and SHINE trials: interpretation of WASH intervention effects on linear growth and diarrhoea.
576 *The Lancet Global Health*. 2019;7(8):e1139-e46.
- 577 20. Gharpure R, Mor SM, Viney M, Hodobo T, Lello J, Siwila J, et al. A One Health Approach to
578 Child Stunting: Evidence and Research Agenda. *The American Journal of Tropical Medicine and*
579 *Hygiene*. 2021;104(5):1620.
- 580 21. Barnes AN, Anderson JD, Mumma J, Mahmud ZH, Cumming O. The association between
581 domestic animal presence and ownership and household drinking water contamination among peri-urban
582 communities of Kisumu, Kenya. *PLoS One*. 2018;13(6):e0197587-e.
- 583 22. Barnes AN, Mumma J, Cumming O. Role, ownership and presence of domestic animals in peri-
584 urban households of Kisumu, Kenya. *Zoonoses Public Health*. 2018;65(1):202-14.
- 585 23. Ngure FM, Humphrey JH, Mbuya MNN, Majo F, Mutasa K, Govha M, et al. Formative research
586 on hygiene behaviors and geophagy among infants and young children and implications of exposure to
587 fecal bacteria. *The American journal of tropical medicine and hygiene*. 2013;89(4):709-16.
- 588 24. Lowenstein C, Waters WF, Roess A, Leibler JH, Graham JP. Animal Husbandry Practices and
589 Perceptions of Zoonotic Infectious Disease Risks Among Livestock Keepers in a Rural Parish of Quito,
590 Ecuador. *The American journal of tropical medicine and hygiene*. 2016;95(6):1450-8.
- 591 25. Thomas ED. Hens, housing, and hygiene: formative research in rural Bangladesh for
592 interventions to separate young children from poultry and poultry feces: Johns Hopkins University; 2021.
- 593 26. Ercumen A, Pickering AJ, Kwong LH, Arnold BF, Parvez SM, Alam M, et al. Animal feces
594 contribute to domestic fecal contamination: evidence from *E. coli* measured in water, hands, food, flies,
595 and soil in Bangladesh. *Environ Sci Technol*. 2017;51(15):8725-34.

- 596 27. Berendes DM, Yang PJ, Lai A, Hu D, Brown J. Estimation of global recoverable human and
597 animal faecal biomass. *Nature Sustainability*. 2018;1(11):679-85.
- 598 28. Ballard AM, Laramée N, Haardörfer R, Freeman MC, Levy K, Caruso BA. Measurement in the
599 study of human exposure to animal feces: A systematic review and audit. *International Journal of Hygiene
600 and Environmental Health*. 2023;249:114146.
- 601 29. Ballard AM, Laramée N, Haardörfer R, Freeman MC, Levy K, Caruso BA. Measurement in the
602 study of human exposure to animal feces database: A systematic review and audit [Internet]. 2022.
603 Available from: osf.io/d2nxc.
- 604 30. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best Practices for
605 Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in
606 Public Health*. 2018;6(149).
- 607 31. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine : A Practical Guide*.
608 Cambridge, UNITED KINGDOM: Cambridge University Press; 2011.
- 609 32. Diamantopoulos A, Winklhofer HM. Index construction with formative indicators: An alternative
610 to scale development. *Journal of marketing research*. 2001;38(2):269-77.
- 611 33. National Research Council. *Exposure science in the 21st century: a vision and a strategy*. 2012.
- 612 34. Lee GO, Eisenberg JNS, Uruchima J, Vasco G, Smith SM, Van Engen A, et al. Gut microbiome,
613 enteric infections and child growth across a rural–urban gradient: protocol for the ECoMiD prospective
614 cohort study. *BMJ Open*. 2021;11(10):e046241.
- 615 35. City Population: Republic of Ecuador, 2010 census 2010 [Available from:
616 <https://www.citypopulation.de/en/ecuador/>].
- 617 36. Smith SM, Montero L, Paez M, Ortega E, Hall E, Bohnert K, et al. Locals get travellers’
618 diarrhoea too: risk factors for diarrhoeal illness and pathogenic *Escherichia coli* infection across an urban-
619 rural gradient in Ecuador. *Tropical Medicine & International Health*. 2019;24(2):205-19.

- 620 37. Lee GO, Whitney HJ, Blum AG, Lybik N, Cevallos W, Trueba G, et al. Household coping
621 strategies associated with unreliable water supplies and diarrhea in Ecuador, an upper-middle-income
622 country. *Water Res.* 2020;170:115269-.
- 623 38. Creswell JW, Creswell JD. *Research design: Qualitative, quantitative, and mixed methods*
624 *approaches*: Sage publications; 2017.
- 625 39. Garcia CM, Eisenberg ME, Frerich EA, Lechner KE, Lust K. Conducting go-along interviews to
626 understand context and promote health. *Qual Health Res.* 2012;22(10):1395-403.
- 627 40. Carpiano RM. Come take a walk with me: the "go-along" interview as a novel method for
628 studying the implications of place for health and well-being. *Health Place.* 2009;15(1):263-72.
- 629 41. Guadagnoli E, Velicer WF. Relation of sample size to the stability of component patterns.
630 *Psychological bulletin.* 1988;103(2):265.
- 631 42. Osborne JW, Costello AB. Sample size and subject to item ratio in principal components
632 analysis. *Practical Assessment, Research, and Evaluation.* 2004;9(1):11.
- 633 43. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments.
634 *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences.*
635 2016;374(2065):20150202.
- 636 44. Kaiser HF. A second generation little jiffy. *Psychometrika.* 1970;35(4):401-15.
- 637 45. Kaiser HF. An index of factorial simplicity. *Psychometrika.* 1974;39:31-6.
- 638 46. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*: Pearson Education; 2013.
- 639 47. BARTLETT MS. THE EFFECT OF STANDARDIZATION ON A χ^2 APPROXIMATION IN
640 FACTOR ANALYSIS. *Biometrika.* 1951;38(3-4):337-44.
- 641 48. Mair P. *Modern psychometrics with R*: Springer; 2018.
- 642 49. Mair PDL, Jan. *Gifi: Multivariate Analysis with Optimal Scaling*. R package version 0.3-10
643 ed2022.
- 644 50. R Team. *RStudio: Integrated Development for R*. Boston, MA: RStudio, PBC; 2020.

- 645 51. Dunteman GH. Principal Components Analysis. Newbury Park, California 1989. Available from:
646 <https://methods.sagepub.com/book/principal-components-analysis>.
- 647 52. Haardörfer R. Taking quantitative data analysis out of the positivist era: Calling for theory-driven
648 data-informed analysis. *Health Education & Behavior*. 2019;46(4):537-40.
- 649 53. Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research. R
650 package version 2.2.5 ed. Northwestern University, Evanston, Illinois 2022.
- 651 54. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for
652 Reliability Research. *Journal of chiropractic medicine*. 2016;15(2):155-63.
- 653 55. Flegal KM, BROWNIE C, HAAS J. The effects of exposure misclassification on estimates of
654 relative risk. *American journal of epidemiology*. 1986;123(4):736-51.
- 655 56. Frongillo EA, Baranowski T, Subar AF, Tooze JA, Kirkpatrick SI. Establishing Validity and
656 Cross-Context Equivalence of Measures and Indicators. *J Acad Nutr Diet*. 2019;119(11):1817-30.
- 657 57. Polites GL, Roberts N, Thatcher J. Conceptualizing models using multidimensional constructs: a
658 review and guidelines for their use. *European Journal of Information Systems*. 2012;21(1):22-48.
- 659 58. Roberts N, Thatcher J. Conceptualizing and testing formative constructs: Tutorial and annotated
660 example. *ACM sigmis database: The database for Advances in Information Systems*. 2009;40(3):9-39.
- 661 59. Ballard AM; Haardörfer R; Corozo Angulo B; Freeman MC; Eisenberg JL; Lee GL; Levy K;
662 Caruso BA. Tools for the development and validation of a survey to measure fecal-oral child exposure to
663 zoonotic enteropathogens: The FECEZ index [Internet]. 2023. Available from: osf.io/jkh3u/

664 **Supporting information**

665 **Table S1. Table of survey items grouped by sub-domain and post-PCA results.**

666 **Table S2. Child sex-disaggregated demographic characteristics (n=297).**

667 **Table S3. Child exposure item frequencies (n=297).**

668 **Fig S1. Correlation plot between child behavior and child environment sub-domain scores (n=297).**

669 **Fig S2. Histogram of child environment sub-domain scores (possible score range: 0-60, n=297).**

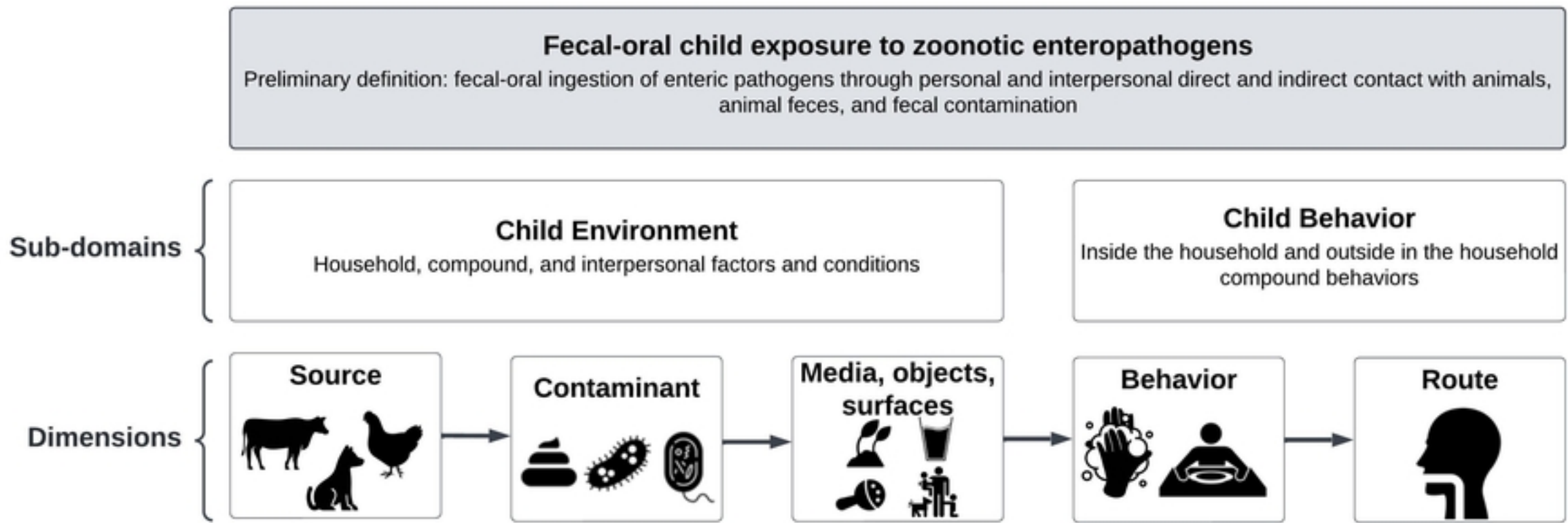


Figure 1

Phase 1: Item Development

Stage 1a: Domain Identification

Purpose: Specify and describe the concept and construct, and develop preliminary conceptual definition

Activities:

- In-depth go-along interviews
- Systematic review

Stage 1b: Item Generation

Purpose: Identify and create items that assess the construct

Activities:

- Inductive and deductive item generation

Stage 1c: Content Validity Assessment

Purpose: Evaluate if items adequately measure the construct of interest

Activities:

- Expert review
- Cognitive interviews

Phase 2: Index Development

Stage 2a: Survey Administration

Purpose: Gather appropriate and sufficient data for index development and evaluation

Activities:

- Establishment of sample frame
- Data collection

Stage 2b: Item Reduction

Purpose: Maximize parsimony of the index

Activities:

- Final sample assessment
- Principal component analysis

Phase 3: Index Evaluation

Stage 3a: Index Item Scoring

Purpose: Create scores for analyses and index evaluation

Activities:

- Calculate component, sub-domain, and overall index scores

Stage 3b: Validity and Reliability Assessment

Purpose: Evaluate repeatability and quality of the index

Activities:

- Construct validity and test-retest reliability evaluation

Figure 2