

1 **Psychometric properties of the Barkin Index of Maternal**  
2 **Functioning with Ethiopian mothers**

3  
4  
5 Selamawit Gebeyehu<sup>1</sup>, Yordanos Gizachew<sup>2</sup>, Yilma Chisa<sup>1</sup>, Beemnt Tekabe<sup>1</sup>, and  
6 Robert Trevethan<sup>3</sup>

7  
8  
9 <sup>1</sup> School of Public Health, College of Medicine and Health Science, Arba Minch University,  
10 Arba Minch, Gamo Zone, Ethiopia

11 <sup>2</sup> School of Nursing, College of Medicine and Health Science, Arba Minch University, Arba  
12 Minch, Gamo Zone, Ethiopia

13 <sup>3</sup> Independent academic researcher and author, Albury, NSW, Australia

14  
15  
16 \*Corresponding Author: Selamawit Gebeyehu

17 Email: emugebe.sg@gmail.com

18  
19 Short title: Psychometric properties of the BIMF with Ethiopian mothers

20

## 21 Abstract

### 22 Background

23 The functional status of mothers after childbirth has implications for maternal, child, and family  
24 health. There is a lack of adequate reliable and valid instruments in Ethiopia for assessing  
25 women's postpartum functional status.

### 26 Objective

27 This study was intended to reveal the psychometric properties of the Barkin Index of Maternal  
28 Functioning (BIMF) for assessing Ethiopian mothers' functional status.

### 29 Method

30 Structured interviews were used to obtain BIMF data from 202 women who had a child less than  
31 1 year of age. Descriptive statistics were calculated for the BIMF items; internal consistency was  
32 assessed with interitem correlations and coefficient alphas; construct validity was examined  
33 through exploratory factor analyses (EFAs) after face and content validities had been confirmed;  
34 and test-retest reliability was assessed with intraclass correlation coefficients (ICCs).

### 35 Results

36 Narrow standard deviations and significant skewness and kurtosis characterized most of the  
37 individual BIMF items. Most interitem correlations were  $< .15$  and 13 of the 20 BIMF items  
38 did not load satisfactorily on any factor in exploratory factor analyses. Two factors emerged  
39 from the remaining items, one with three items and the other with four. Coefficient alphas were

40 .54 for the first of these factors, .48 for the second, and .58 for all 20 items. The ICCs for test–  
41 retest reliability were < .40.

## 42 **Conclusions**

43 BIMF data from the sample of Ethiopian women in this study exhibited unusually low levels of  
44 variability and high levels of skewness and kurtosis. Furthermore, in the context of this study, the  
45 BIMF could not be regarded as reliable either in terms of internal (interitem) consistency or  
46 temporal (test–retest) consistency. Researchers using the BIMF in Ethiopian contexts are advised  
47 to examine the nature of their data carefully, identify the factor structure of their samples’ data,  
48 and consider ways in which the index might be improved, or replaced.

49

50 **Keywords:** Barkin Index of Maternal Functioning, BIMF, maternal functioning, Ethiopia

51

## 52 **Introduction**

53 Maternal functioning in the postpartum period is a reflection of the physical and psychosocial  
54 changes associated with pregnancy and childbirth. Maternal activities, in addition to  
55 physiological changes as well as individual and social activities, can make this period difficult  
56 for mothers. As a result, assessing maternal functioning following childbirth can be important for  
57 identifying maternal performance, productivity, and health, and can provide valuable information  
58 for healthcare providers, particularly with regard to identifying mothers who might benefit from  
59 supportive advice or interventions [1–3].

60  
61 Almost all of the postpartum functional status assessment tools available today were  
62 developed for populations in high-income countries [4]. To date, the Inventory of Functional  
63 Status After Childbirth (IFSAC) is the most widely used tool specifically designed to assess  
64 several domains of maternal functioning [5]. The IFSAC contains 36 items arranged in subscales  
65 encompassing five dimensions: infant care responsibilities, self-care activities, household  
66 activities, social and community activities, and occupational activities. The inventory is designed  
67 to assess maternal functioning from 6 to 10 weeks after childbirth with a focus on “recovery  
68 from childbirth” [6]. Despite its popularity, the IFSAC has been criticized because it does not  
69 account for women’s feelings or levels of satisfaction with the changes in their lives since  
70 childbirth, and, as an alternative, the Barkin Index of Maternal Functioning (BIMF) was  
71 recommended by its developers as a more modern, patient-centered method for assessing  
72 maternal functioning [7, 8]. It is therefore the focus of this study.

73

74 The BIMF is a 20-item Likert-type scale designed to capture mothers' perceptions of their  
75 functional status across the previous 2-week period at some point within the year since the birth  
76 of their most-recent, or only, child. For each item, participants are offered six response options: 0  
77 corresponding to "strongly disagree", 1 to "disagree", 2 to "slightly disagree", 3 to "neutral", 4 to  
78 "slightly agree", 5 to "agree", and 6 to "strongly agree". Items 16 and 18 are negatively worded,  
79 so responses on those items are reverse coded prior to analyzing the data to ensure that higher  
80 scores on the index consistently indicate higher levels of functioning.

81

82 Because of the lack of a valid means for assessing maternal functional status in sub-Saharan  
83 countries, including Ethiopia, in this article we examine the validity and reliability of the BIMF  
84 when used with Ethiopian mothers to identify whether the index is likely to perform  
85 satisfactorily within that population.

86

## 87 **Method**

### 88 **Study design and setting**

89 A descriptive community-based study was implemented with participants from Lante District  
90 within Gamo Zone, Ethiopia, situated 430 km south of the country's capital, Addis Ababa. In  
91 2020 G.C. (2012 E.C.), the district health office reported a total population of 236,600, with 436  
92 mothers having a child under 1 year of age.

93

### 94 **Participants**

95 Eligible participants were women who had a child less than 1 year of age resulting from a single-  
96 child pregnancy; who had vaginal delivery, birth through caesarean section, or instrumental

97 delivery; who had self-declared physical health; and who had self-declared absence of depression  
98 or any other mental disorder during the pre-pregnancy, prepartum, and postpartum periods.

99

## 100 **Sample size determination and sampling method**

101 Sample size required to determine the factor structure of the BIMF was based on there being at  
102 least 10 participants per item of the index [9]. Therefore, a sample of at least 200 mothers was  
103 required. In order to attain that sample, a list of the 436 mothers who had a child less than 1 year  
104 of age was obtained from the district health office. Subsequently, a sample of 202 mothers was  
105 identified by using computer-generated random numbers with Microsoft Excel for Windows 10.  
106 Forty of the 202 mothers were selected by the same method to comprise a retest sample for the  
107 BIMF, 2 weeks after the index had initially been administered.

108

## 109 **Questionnaire**

110 A questionnaire was created that comprised questions to obtain information about the mothers'  
111 sociodemographic background (the respondent and her most-recent child's ages, the  
112 respondent's marital status, and size of her family), obstetric characteristics (parity, number of  
113 living children, mode of delivery of the most recent child, and whether or not the mother had  
114 intended to become pregnant), and the BIMF.

115

## 116 **Translation process**

117 Translation and cultural adaptation of the BIMF (for example, the word "diaper" is not familiar  
118 to most people in rural communities, and in Amharic is referred to as "ye shint cherq") was  
119 performed according to the minimal translation criteria of the Medical Outcomes Trust [10]. Two

120 independent bilingual translators (one a psychologist and the other an English-language lecturer)  
121 with advanced levels of English language and native Amharic-language skills translated the  
122 BIMF into Amharic. In doing so, they made no changes to the items' meanings. An English-  
123 language lecturer who is a native Amharic speaker and who had been blinded to the original  
124 version, provided a back translation. There were no major difficulties in reconciling the back-  
125 translated version. Subsequently, Dr Barkin, the main researcher associated with development of  
126 the BIMF, gave permission for the Amharic version of the index to be used in this research.

127

### 128 **Face and content validity**

129 After translation, the BIMF was reviewed by experts in maternity and child health and  
130 reproductive health units who were not part of the study, and they identified no problems. The  
131 tool was regarded as likely to be generally well understood, acceptable, and culturally  
132 appropriate for all prospective respondents.

133

### 134 **Data collection**

135 Names and addresses of the 202 mothers identified as prospective participants were obtained  
136 from the district health office. Five data collectors and three field supervisors who had  
137 experience in conducting interviewer-administered data-collection procedures and the ability to  
138 speak Amharic were recruited for data collection. The data collectors went to the mothers'  
139 homes and conducted structured interviews by using the smart-phone based Open Data Kit  
140 (ODK) app. Prospective participants were provided verbal and written information about the  
141 research, and were invited to ask any questions they might have. All mothers who were invited to  
142 participate in the research agreed to do so.

143

## 144 **Data processing and analysis**

145 Data were analyzed using SPSS for Windows version 25<sup>®</sup> (IBM Corp., Armonk, NY, USA).

146 Descriptive statistics were calculated for the sociodemographic variables and BIMF items.

147 Reliability associated with internal consistency was measured by interitem correlations and

148 coefficient alphas.

149 Because previous research with the BIMF had revealed marked inconsistency regarding number

150 and composition of factors in the scale [11–14], we did not conduct confirmatory factor analysis

151 as a means of identifying the structure of the data. Instead, we conducted a series of exploratory

152 factor analyses (EFAs). In doing so, we identified the number of factors with the scree test in

153 conjunction with parallel analysis based on principal components and 1,000 randomly generated

154 matrices; we used principal axis factoring as the method of extraction in conjunction with

155 promax rotations; and we regarded item loadings  $> |0.40|$  as definitely acceptable, and loadings

156 in the region of  $|.35|$  to  $|.40|$  as possibly acceptable.

157

158 In order to assess test–retest reliability, ICCs (3,1 [two-way mixed effects, single measures]

159 absolute agreement) were calculated.

160

## 161 **Ethical considerations**

162 Ethics approval was obtained from Arba Minch University’s College of Medicine and Health

163 Science Institutional Review Board (IRB/157/2020). Verbal consent was obtained from all

164 participants and recorded in conjunction with the starting time of each interview, and all

165 interviews were conducted in private.



166 **Results**

167 **Participant characteristics**

168 All 202 mothers reported their physical health to be good, and none indicated they had a history  
169 of depression or any other mental disorder during the pre-pregnancy, prepartum, and postpartum  
170 periods. Their ages ranged from 15 to 38 years ( $M = 26.1$  years, mode = 30,  $SD = 4.7$ ), their  
171 children's ages ranged from 1 to 12 months ( $M = 6.7$  months, mode = 11,  $SD = 3.7$ ), their  
172 number of pregnancies ranged from 1 to 9 ( $M = 2.4$ , mode = 1,  $SD = 1.5$ ), their number of  
173 children ranged from 1 to 8 ( $M = 2.31$ , mode = 1,  $SD = 1.40$ ), and the size of their households  
174 ranged from 2 to 12 ( $M = 4.6$ , mode = 4,  $SD = 1.7$ ). Other participant characteristics are shown in  
175 Table 1.

176 **Table 1. Participant characteristics.**

<b>Variable</b>	<b>Category</b>	<b>Number</b>	<b>Percent</b>
<b>Age of the mother/ caregiver (in years)</b>	≤ 24	73	36.1
	25 – 30	103	51.0
	≥ 31	26	12.9
<b>Age of most recent or only child</b>	< 2 months	20	9.9
	2 – 6 months	76	37.6
	≥ 6 months	106	52.6
<b>Marital status</b>	Single	9	4.5
	Married	191	94.6
	Divorced or widowed	2	1
<b>Family size</b>	≤ 3	58	28.7
	4 – 6	115	56.9
	≥ 7	29	14.4

<b>Parity</b>	Primipara	69	34.2
	Multipara	133	65.8
<b>Number of living children</b>	≤ 3	165	81.7
	≥ 4	37	18.3
<b>Mode of delivery</b>	Vaginal	164	81.2
	Caesarean	25	12.4
	Instrument	13	6.4
<b>Pregnancy intention</b>	Intended	175	86.6
	Unintended	27	13.4

177

178

## 179 **Item characteristics**

180 Means, *SDs*, and indicators of skewness and kurtosis on the 20 BIMF items at Time 1 are  
 181 provided in Table 2. On all but one of the items, the means were above the midpoint of the  
 182 option range, indicating a tendency among participants to believe they were functioning  
 183 effectively as mothers. The exception was Item 18, on which, prior to reverse coding, the  
 184 participants indicated they were often anxious or worried about fulfilling their maternal  
 185 responsibilities. Twelve of the items had *SDs* < 1.0; 18 items were significantly skewed (two  
 186 positively, 16 negatively); and 15 items were significantly kurtotic (12 leptokurtic, three  
 187 platykurtic). Twelve items were significantly negatively skewed as well as being significantly  
 188 leptokurtic.

189

190 **Table 2. Descriptive statistics for items on the Barkin Index of Maternal Functioning.**

Item <sup>a</sup>	Mean	<i>SD</i>	Skew- ness <i>z</i> <sup>b</sup>	Kurto- sis <i>z</i> <sup>b</sup>
-------------------	------	-----------	--	--

1	I am a good mother.	5.2	.46	4.3	0.2
2	I feel rested.	5.0	.91	-16.2	33.7
3	I am comfortable with the way I've chosen to feed my baby (either bottle or breast or both).	5.0	.55	-15.1	54.4
4	My baby and I understand each other.	5.2	.58	-0.2	-0.7
5	I am able to relax and enjoy time with my baby.	5.4	.69	-5.6	4.7
6	There are people in my life that I can trust to take care for my baby when I need a break.	4.6	1.25	-12.0	10.6
7	I am comfortable allowing a trusted friend or relative to care for my child (can include baby's father or partner).	4.7	1.29	-11.9	10.1
8	I am getting enough adult interaction.	4.9	.61	-7.4	22.4
9	I am getting enough encouragement from other people.	4.7	.80	-13.4	26.4
10	I trust my own feelings (instincts) when it comes to taking care of my baby.	3.5	2.00	-3.9	-3.8
11	I take a little time each week to do something for myself.	3.1	1.84	-1.3	-4.9
12	I am taking good care of my baby's physical needs (feeding, "ye shint cherq", and keeping doctor's appointments).	5.2	.80	-15.4	40.7
13	I am taking good care of my physical needs (eating, showering, etc.).	5.0	.79	-7.8	13.3
14	I make good decisions about my baby's health and well being.	4.9	.71	-11.1	26.7
15	My baby and I are getting into a routine.	3.6	1.83	-3.4	-4.1
16	I [do not] worry about how other people judge me (as a mother).	4.2	1.70	-7.2	-0.2
17	I am able to take care of my baby and my other responsibilities.	4.8	.96	-14.8	25.7
18	Anxiety or worry [seldom interfere] with my mothering ability.	2.1	1.73	6.3	-1.4
19	As time goes on, I am getting better at taking care of my baby.	4.8	1.03	-11.7	19.7
20	I am satisfied with the job I am doing as a new mother.	5.5	.63	-4.3	-1.3

191 <sup>a</sup> Responses to Items 16 and 18 have been reverse coded to attain the meaning created by bracketed text.

192 <sup>b</sup> The  $z$  values for skewness and kurtosis are based on original skewness and kurtosis values divided by  
193 their respective standard errors. If these  $z$  values exceed  $|3.29|$  when  $N = 50$  to  $300$ , they indicate  
194 statistically significant skewness or kurtosis [15].

195

## 196 **Interitem correlations and factor structure**

197 None of the 190 interitem correlations exceeded  $.50$ , and only 37% of those correlations lay in  
198 the  $.15$  to  $.50$  range recommended by Clark and Watson [16]. Among the remaining 63% of the  
199 interitem correlations, 17% were negative, with the strongest being  $-.29$ . Despite this generally  
200 low, and sometimes inverse, level of association between the items, the Kaiser–Meyer–Olkin  
201 (KMO) index was  $.72$  and Bartlett’s test of sphericity was significant ( $p < .001$ ), and therefore  
202 the data appeared to be suitable for EFA. Among the 20 items, seven eigenvalues were  $> 1$ , but  
203 the scree plot indicated there were four factors in the data (see Fig 1), as did parallel analysis. We  
204 therefore conducted an initial EFA constrained to four factors. In this solution, nine items loaded  
205 uniquely on one of the factors at  $> |.40|$ , three items had cross loadings on two of the factors, and  
206 the remaining eight items failed to load on any factor at  $> |.37|$ . Only 42.56% of the variance was  
207 accounted for by the four factors.

208

209 [Fig1. Scree plot based on all 20 BIMF items.]

210

211 Follow-up EFAs were conducted in order to attain the most satisfactory and interpretable  
212 outcome. When nonloading items, cross-loading items, and items that loaded only on single-item  
213 factors were successively removed, seven items were retained, among which 53.09% of the  
214 variance was accounted for, but the extraction communalities were generally low (maximum =  
215  $.55$ , mean =  $.35$ , with Items 20 and 18 exceptionally low at  $.23$  and  $.13$ , respectively). All seven

216 items loaded at  $> .39$  on one of two factors, although two items cross loaded with a difference  $\leq$   
217  $.20$ . See Table 3. Items 13, 14, and 17 loaded on a factor that we labeled Decision Making and  
218 Practicalities, and Items 5, 16, 18, and 20 loaded on the other factor, which we labeled  
219 Confidence, Enjoyment, and Satisfaction.

220

221 **Table 3. Item loadings on the two-factor solution.<sup>a</sup>**

Item #	Item <sup>b</sup>	Factor 1	Factor 2
14	I make good decisions about my baby's health and well-being.	.74	
17	I am able to take care of my baby and my other responsibilities.	.56	-.40
13	I am taking good care of my physical needs (eating, showering, etc.).	.52	
16	I [do not] worry about how other people judge me (as a mother).		.64
5	I am able to relax and enjoy time with my baby.	.32	.52
20	I am satisfied with the job I am doing as a new mother.		.42
18	Anxiety or worry [seldom interfere] with my mothering ability.		.39

222

223 <sup>a</sup> Loadings  $< |.30|$  are not shown.

224 <sup>b</sup> Wording of Items 16 and 18 was altered, as shown in bracketed text, to reflect prior reverse coding.

225

226 Consonant with the practice of several other researchers [11, 13, 14], we calculated a score for  
227 each subscale by adding responses on their respective items, and we calculated a composite total  
228 by adding responses to all 20 items. In order to confer meaning to participants' responses in  
229 relation to the range of response options, we also divided these totals by the number of items on  
230 which each score was based. The results are shown in Table 4.

231

232 **Table 4. Means, standard deviations, and obtainable ranges on composite factor and total**  
233 **scores.<sup>a</sup>**

BIMF <sup>b</sup> component and items	Unadjusted scores		Adjusted scores	
	Mean ( <i>SD</i> )	Obtainable range	Mean ( <i>SD</i> )	Obtainable range
Decision Making and Practicalities (Items 13, 14, 17)	14.74 (1.79)	0–18	4.91 (0.60)	0–6
Confidence, Enjoyment, and Satisfaction (Items 5, 16, 18, 20)	17.08 (1.79)	0–24	4.27 (0.81)	0–6
Total (Items 1 to 20)	91.11 (7.76)	0–120	4.56 (0.39)	0–6

234  
235 <sup>a</sup> Unadjusted scores are based on addition of mothers' responses; adjusted scores were obtained by  
236 dividing the unadjusted scores by the relevant number of items.

237 <sup>b</sup> Barkin Index of Maternal Functioning

238  
239 The means of the adjusted scores on both factors as well as the composite total indicate that  
240 the mothers tended to respond with options of somewhat agree (4) or agree (5). However,  
241 relative to the first factor (Decision Making and Practicalities), the means and *SDs* indicated that  
242 self-perceptions of maternal functioning were lower and more varied on the second factor  
243 (Confidence, Enjoyment, and Satisfaction). Scores on the first factor were significantly skewed  
244 (–11.76) and kurtotic (21.9), and the total score was significantly skewed (–3.4).

245  
246 Coefficient alphas were .54 for Decision Making and Practicalities, .48 for Confidence,  
247 Enjoyment, and Satisfaction, and .58 for all 20 items.

## 248 **Test–retest results**

249

250 The ICC for Confidence, Enjoyment, and Satisfaction could not be calculated satisfactorily  
251 because of negative average covariance among the items (Pearson's *r* of –.04 was calculated as a

252 substitute for the ICC). The ICC for Decision Making and Practicalities was .14 [95% CI -.18–  
253 .43], and the ICC for the total score across all 20 items was .36 [95% CI .06–.61].

254

## 255 **Discussion**

256 To the authors' best knowledge, this is the first examination of the BIMF in an Ethiopian  
257 context. The research has several desirable foundations. Face and content validity of the BIMF  
258 were assessed and regarded as satisfactory; random sampling was used to obtain participants; the  
259 sample included women who gave birth by vaginal delivery, instrument, and caesarean section,  
260 and who were primiparous as well as multiparous, suggesting that the results were not likely to  
261 be unique to a specific group of mothers; the sample was sufficiently large for identifying the  
262 factor structure of the data; and an attempt was made to identify test–retest reliability.

263

264 Three aspects of the raw data are noteworthy and might carry implications for assessing the  
265 suitability of the BIMF in Ethiopian settings. First, with the exception of Item 18 (referring to  
266 mothers' worry or anxiety), the participants' responses on individual items suggest a moderate to  
267 high level of maternal functioning. Whether that is an accurate indication would require  
268 independent assessment, which was not conducted in this research. However, some reassurance  
269 might be taken from the mean of 91 on the 20 BIMF items in this research being similar to the  
270 mean of the 20 items in research where samples were composed solely of mothers who were not  
271 depressed [13, 17]. Furthermore, in samples that include a large proportion of mothers diagnosed  
272 with depression and who could therefore be expected to have lower 20-item BIMF scores, the  
273 mean based on those items has indeed been lower, for example, ranging from 81 down to 63 [7,  
274 11, 17, 18]. However, in a sample of mothers, 46% of whom had been assessed as experiencing

275 depression, the 20-item BIMF score was considerably higher, at 104 [19]. Unfortunately, this  
276 range of 40 points on the total BIMF score in samples with mothers who experience depression  
277 makes it difficult to place confidence in the meaning of the 20-item BIMF score when seeking to  
278 gain a definitive overview of mothers' functioning, and the composite score of 91 in this research  
279 therefore provides little reassurance in relation to validity of the BIMF in an Ethiopian setting.

280  
281 The second noteworthy aspect of the data is the lack of variability among the mothers'  
282 responses. This was evident in 60% of the *SDs* on individual items being < 1.00 despite the  
283 availability of seven response options. In contrast, only two of the *SDs* were < 1 on the 20 items  
284 in the foundational psychometric research with the BIMF [11]. Furthermore, the *SD* of the  
285 composite score based on all 20 items in this research was 7.76 compared with noticeably wider  
286 *SDs*, ranging from 10.8 in some research [20] up to 17.1 in other research [7].

287  
288 This smaller amount of variation in the mothers' responses in the present study presents  
289 interpretative and empirical difficulties for three reasons. First, if there is genuine lack of  
290 variability, statistics such as correlations, factor analysis, and coefficient alpha—all of which rely  
291 on variability—are likely to yield misleading results. Second, if the lack of variability has a  
292 foundation in some kind of distorting BIMF-related artifact within Ethiopian samples (e.g., the  
293 cognitive burden on the mothers having to remember what occurred during the previous 2 weeks  
294 might have created a degree of bluntness in their responses), the nature of that artifact would  
295 need to be identified and addressed. Third, if the lack of variability is likely to be a chance  
296 occurrence, other studies need to be conducted with the prospect that their data would contain a  
297 greater spread of maternal self-impressions.



298

299       The third noteworthy aspect of the data is most items having distributions that were  
300 significantly skewed or kurtotic, or both, which was accompanied by one of the factors and the  
301 20-item composite score having negatively skewed distributions. These results could be genuine  
302 reflections of the mothers' self-perceptions, and therefore are interesting findings in themselves.  
303 However, in combination with the minimal variation in responses to the 20 items, the skewed  
304 and kurtotic distributions might also have influenced the EFA outcome, coefficient alpha values,  
305 and the test–retest results.

306

307       We are aware of no other research in which a similar BIMF factor structure emerged.  
308 Although two factors have been found in some other studies [11, 14], the factors in those studies  
309 comprised different items and were therefore conceived of as being different and were labeled  
310 differently. Our results are also distinctive in that 13 of the 20 BIMF items failed to load on any  
311 factor, probably because most interitem correlations were  $< .15$ . An additional feature of our  
312 results is Items 16 and 18 having loaded on one of our factors (Confidence, Enjoyment, and  
313 Satisfaction), but having loaded on none of the factors in some other research [11, 12, 14].

314

315       Although inconsistencies in factor structure could result from differences in the nature of  
316 samples, particularly when cultural differences exist, some inconsistency between our findings  
317 and the findings in other research is likely to result from principal components analysis versus  
318 EFA being employed as a data-reduction strategy; different procedures (the Kaiser criterion,  
319 scree plots, or parallel analysis) being used for identifying the number of factors in the data; the  
320 method used for extraction (e.g., maximum likelihood or principal axis factoring); whether or not

321 the factors were rotated; the method employed for rotation (orthogonal versus oblique)—if factor  
322 rotation occurred at all; and whether or not a sequence of EFAs was used to attain greater factor  
323 clarity by successively deleting poorly performing items. Publications about the BIMF provide  
324 little or no information about these procedures when data-reduction analyses have been  
325 conducted despite some procedures having substantial implications for the outcomes.

326  
327 In this research, we believe we engaged in best-practice data-reduction procedures [21–26], so  
328 the results are likely to have been dependent on the nature of the data, not deficiencies in the  
329 method of analysis. That aside, the EFA results in this research are not impressive. Not only  
330 were loadings on 13 of the 20 BIMF items so low that most of the index's items were discarded,  
331 but also the remaining seven items had low communalities despite accounting for 53.1% of the  
332 variance. Furthermore, two of the seven items cross loaded on the two factors that emerged.  
333 Under these circumstances, one solution involves focusing on the composite score based on all  
334 20 items, but mothers with similar scores could have noticeably different profiles and therefore  
335 noticeably different needs in relation to healthcare as well as correlates in research contexts, so  
336 the value of that score is limited.

337  
338 In addition to differences in factor structure, our results differ in several ways from the results  
339 obtained in other studies. The coefficient alpha across all 20 items in our data was unusually low  
340 at .58 compared with alphas of .87 and .83 in samples of American mothers [7, 11, 19] and, at  
341 .36, the test–retest ICC for the 20-item composite score in this research indicated much poorer  
342 test–retest reliability than was indicated by the ICC of .80 obtained in a study conducted with  
343 Iranian mothers [14]. The above differences, in which the psychometric features of the BIMF in

344 this research are not only less impressive than are results reported in other BIMF research but are  
345 also unimpressive according to conventional criteria, are almost definitely attributable to limited  
346 data variability in this study [27, 28].

347

348 Although not directly associated with our examination of the BIMF, the absence of self-  
349 reported depression among the mothers in this research might be important. Other research has  
350 indicated that perinatal depression is likely to be approximately 13% among mothers in middle-  
351 and low-income countries in general [29] and in Ethiopia, specifically [30], and as high as 33%  
352 in rural parts of Ethiopia [31]. The discrepancy between those statistics and the mothers' self-  
353 reports in this research suggests that responses on the BIMF in this research might be regarded  
354 with a degree of reservation despite the mothers having been assured by the data collectors of  
355 confidentiality concerning their responses in order to discourage them from creating an  
356 inaccurately favorable impression.

357

## 358 **Conclusions**

359 This study provides ambiguous evidence concerning the reliability and validity of the BIMF in  
360 an Ethiopian setting, largely because it is difficult to identify whether the generally high and  
361 similar self-perceptions of mothers were genuine and might also have been idiosyncratic to our  
362 specific sample.

363

364 The ambiguity could lead to incorrect conclusions that, in Ethiopian samples, the BIMF's  
365 items carry little association with each other; that there are only two factors within the BIMF,  
366 each with only a small number of items; and that the index's test-retest reliability is poor.

367 Because these conclusions might be proved untenable if data were obtained from other samples,  
368 including samples with a greater dispersion among responses, the BIMF may be appropriate in  
369 other Ethiopian contexts. Furthermore, with more variation in the data there is a prospect that  
370 additional, or more substantial, factors would emerge.

371

372 Because the two factors that emerged from the EFAs were dissimilar to factors found in other  
373 research, and because of evident volatility in the factor structure of the BIMF revealed by other  
374 studies, further use of the BIMF in an Ethiopian context should involve examination of the factor  
375 structure for the particular sample(s) on which that research is conducted. Furthermore, because  
376 our research indicates that the index could not be regarded as reliable either in terms of internal  
377 (interitem) consistency or temporal (test–retest) consistency, prospects for improving the index  
378 might be considered—although with the proviso that permission to modify the index had been  
379 granted [32].

380

381 The seven BIMF items that loaded on two factors in this research might be used to explore  
382 aspects of decision making and practicalities, as well as aspects of enjoyment and satisfaction,  
383 among Ethiopian women concerning their maternal role in the postpartum period. However,  
384 given the volatility of the BIMF factor structure, those two factors might not feature in data from  
385 other samples. Furthermore, if only those seven items were used, important aspects of  
386 motherhood such as mother–infant interaction, self-care, and social support would not be  
387 assessed, and this raises the possibility that, when face and content validity were assessed at the  
388 start of this research, too much consideration was given to endorsing the items within the BIMF

389 without considering aspects of motherhood that were underrepresented, inappropriately  
390 represented, or not represented at all — at least for an Ethiopian context.

391  
392 In summary, this research demonstrates that the BIMF should be used with a degree of  
393 caution in Ethiopian contexts and that, if researchers want to use it, they should anticipate  
394 obtaining sufficiently large samples to be able to conduct EFAs to identify the factor structure of  
395 the data produced by those samples and also consider ways in which the index might be modified  
396 if doing so is likely to result in more varied responses — with the underlying proviso that the  
397 greater variability was associated with genuine differences in self-perception. An alternative  
398 course of action would be to seek, or create, instruments that provide greater applicability and  
399 usefulness for assessing maternal functioning in Ethiopian contexts.

400

## 401 **Acknowledgements**

402 The authors thank the participants for their valuable information and the data collectors and  
403 supervisors for their commitment to this research. We also extend our gratitude to Arba Minch  
404 University for support and opportunities provided for us to conduct this study.

405

## 406 **Abbreviations and acronyms**

407 BIMF: Barkin Index of Maternal Functioning

408 EFA: exploratory factor analysis

409 ICC: intraclass correlation coefficient

410 IFSAC: Inventory of Functional Status After Childbirth

## 411 **References**

- 412 1. Aydin J, Barkin JL, Kukulcu K. Attempts to strengthen maternal functioning in the  
413 postpartum period: a literature review. *J Hum Sci.* 2016; 13: 5751–5759.  
414 <https://doi.org/10.14687/jhs.v13i3.4170>
- 415 2. Ahn S, Youngblut JM. Predictors of women’s postpartum health status in the first 3  
416 months. *Asian Nurs Res.* 2007; 1: 136–146. [https://doi.org/10.1016/S1976-  
417 1317\(08\)60016-X](https://doi.org/10.1016/S1976-1317(08)60016-X)
- 418 3. Dulaney KJ, Hallwood DM, Keeter KZ. Post-pregnancy functional ability. Master’s  
419 Thesis, Grand Valley State University Allendale, Michigan. 1998. Available from:  
420 <https://scholarworks.gvsu.edu/theses/415>
- 421 4. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a  
422 questionnaire in perioperative and pain medicine. *Saudi J Anaesth.* 2017; 11: S80–S89.  
423 [https://doi.org/10.4103/sja.SJA\\_203\\_17](https://doi.org/10.4103/sja.SJA_203_17)
- 424 5. Sultan P, Sadana N, Sharawi N, Blake L, El-Boghdadly K, Falvo A, et al. Evaluation of  
425 domains of patient-reported outcome measures for recovery after childbirth: a scoping and  
426 systematic review. *JAMA Netw Open.* 2020; 3: 1–12.  
427 <https://doi.org/10.1001/jamanetworkopen.2020.5540>
- 428 6. Fawcett J, Tulman L, Myers ST. Development of the Inventory of Functional Status  
429 After Childbirth. *J Nurse Midwifery.* 1988; 33: 252–260. [https://doi.org/10.1016/0091-  
430 2182\(88\)90080-8](https://doi.org/10.1016/0091-2182(88)90080-8)
- 431 7. Barkin JL, Wisner, KL, Bromberger, JT, Beach, SR, Terry, MA, Wisniewski, SR.  
432 Development of the Barkin Index of Maternal Functioning. *J Womens Health.* 2010;  
433 19(12): 2239–2246. <https://doi.org/10.1089/jwh.2009.1893>

- 434 8. Barkin JL, Wisner KL, Bromberger JT, Beach SR, Wisniewski SR. Assessment of  
435 functioning in new mothers. *J Womens Health*. 2010; 19: 1493–1499.  
436 <https://doi.org/10.1089/jwh.2009.1864>
- 437 9. Dixon, AE. Exploratory factor analysis. In Plitchta SB, Kelvin EA, editors. *Munro’s*  
438 *statistical methods for health care research* (6th ed.). Philadelphia, PA. Wolters Kluwer;  
439 2013. pp. 371–398.
- 440 10. Medical Outcomes Trust. Trust introduces new translation criteria. *Med Outcomes Trust*  
441 *Bull*. 1997; 5: 3–4.
- 442 11. Barkin JL, Wisner KL, Wisniewski SR. The psychometric properties of the Barkin Index  
443 of Maternal Functioning. *J Obstet Gynecol Neonatal Nurs*. 2014; 43: 792–802.  
444 <https://doi.org/10.1111/1552-6909.12505>
- 445 12. Aydın R, Kukulu K. Adaptation of the Barkin Scale of Maternal Functioning and  
446 examination of the psychometric properties. *Health Care Women Int*. 2018; 39: 50–  
447 64. <https://doi.org/10.1080/07399332.2017.1385616>
- 448 13. Ansariniaki M, Lamyian M, Foroushani AR, Curry CL, Barkin JL. Persian version of the  
449 Barkin Index of Maternal Functioning (BIMF): a cross-cultural adaptation and  
450 psychometric evaluation. *BMC Pregnancy Childbirth*. 2020; 21: 83.  
451 <https://doi.org/10.1186/s12884-021-03556-4>
- 452 14. Mirghafourvand M, Barkin JL, Jafarabadi MA, Karami F, Ghanbari-Homayi S. The  
453 psychometric properties of the Barkin Index of Maternal Functioning (BIMF) for the  
454 Iranian population. *BMC Womens Health*. 2019; 19: 166. [https://doi.org/10.1186/s12905-](https://doi.org/10.1186/s12905-019-0859-2)  
455 [019-0859-2](https://doi.org/10.1186/s12905-019-0859-2)

- 456 15. Kim H-Y. Statistical notes for clinical researchers: assessing normal distribution (2) using  
457 skewness and kurtosis. *Restor Dent Endod*. 2013; 38: 52–54.  
458 <https://doi.org/10.5395/rde.2013.38.1.52>
- 459 16. Clark LA, Watson D. Constructing validity: basic issues in objective scale  
460 development. *Psychol Assess*. 1995; 7: 309–319. [https://doi.org/10.1037/1040-](https://doi.org/10.1037/1040-3590.7.3.309)  
461 [3590.7.3.309](https://doi.org/10.1037/1040-3590.7.3.309)
- 462 17. Chamgurdani FK, Barkin JL, Esmailpour K, Malakouti J, Buoli M, Mirghafourvand M.  
463 The effect of counseling with a skills training approach on maternal functioning: a  
464 randomized controlled clinical trial. *BMC Womens Health*. 2020; 20: 51.  
465 <https://doi.org/10.1186/s12905-020-00914-w>
- 466 18. Barkin JL, Wisner KL, Bromberger JT, Beach SR, Wisniewski SR. Factors associated  
467 with postpartum maternal functioning. *J Womens Health*. 2016; 25: 707–713.  
468 <https://doi.org/10.1089/jwh.2015.5296>
- 469 19. Barkin JL, McKeever A, Lian B, Wisniewski SR. Correlates of postpartum maternal  
470 functioning in a low-income obstetric population. *J Am Psychiatr Nurses Assoc*.  
471 2017; 23: 149–158. <https://doi.org/10.1177/1078390317696783>
- 472 20. Barkin JL, Willis GB, Hawkins KC, Stanfill-Thomas T, Beals L, Bloch JR. Semantic  
473 assessment of the Barkin Index of Maternal Functioning in a medically underserved  
474 obstetric population. *Perspect Psychiatr Care*. 2017; 53: 95–103.  
475 <https://doi.org/10.1111/ppc.12141>
- 476 21. Beavers AS, Lounsbury JW, Richards JK, Huck SW, Skolits GJ, Esquivel SL. Practical  
477 considerations for using exploratory factor analysis in educational research. *Practical*  
478 *Assess Res Evaluation*. 2013; 18: 6. <https://doi.org/10.7275/qv2q-rk76>



- 479 22. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four  
480 recommendations for getting the most from your analysis. *Practical Assess Res Evaluation*.  
481 2005; 10: 7. <https://doi.org/10.7275/jyj1-4868>
- 482 23. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory  
483 factor analysis in psychological research. *Psychol Methods*. 1999; 4(3): 272–299.  
484 <https://doi.org/10.1037/1082-989X.4.3.272>
- 485 24. Gaskin CJ, Happell B. On exploratory factor analysis: a review of recent evidence, an  
486 assessment of current practice, and recommendations for future use. *Int J Nurs Stud*. 2014;  
487 51(3): 511–521. <https://doi.org/10.1016/j.ijnurstu.2013.10.005>
- 488 25. Matsunaga M. How to factor-analyze your data right: do's, don'ts, and how-to's. *Int J*  
489 *Psychol Res*. 2010; 3(1): 97–110. <https://doi.org/10.21500/20112084.854>
- 490 26. Preacher KJ, MacCallum RC. Repairing Tom Swift's electric factor analysis machine.  
491 *Underst Statistics*. 2003; 2(1): 13–43. [https://doi.org/10.1207/S15328031US0201\\_02](https://doi.org/10.1207/S15328031US0201_02)
- 492 27. Goodwin LD, Leech NL. Understanding correlation: Factors that affect the size of r. *J*  
493 *Exp Educ*. 2006; 74: 249–266. <https://doi.org/10.3200/JEXE.74.3.249-266>
- 494 28. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation  
495 coefficients for reliability research. *J Chiropr Med*. 2016; 15: 155–163.  
496 <https://doi.org/10.1016/j.jcm.2016.02.012>
- 497 29. Woody CA, Ferrari AJ, Siskind DJ, Whiteford HA, Harris MG. A systematic review and  
498 meta-regression of the prevalence and incidence of perinatal depression. *J Affect Disord*.  
499 2017; 219: 86–92. <https://doi.org/10.1016/j.jad.2017.05.003>

- 500 30. Molenaar J, Hanlon C., Alem A, Wondimagegn D, Medhin G, Prince M, Stevenson  
501 EGJ. Perinatal mental distress in a rural Ethiopian community: a critical examination of  
502 psychiatric labels. *BMC Psychiatry*. 2020; 20: 223. [https://doi.org/10.1186/s12888-020-](https://doi.org/10.1186/s12888-020-02646-5)  
503 02646-5
- 504 31. Kerie S, Menberu M, Niguse W. Prevalence and associated factors of postpartum  
505 depression in Southwest, Ethiopia, 2017: a cross-sectional study. *BMC Res Notes*.  
506 2018; 11: 623. <https://doi.org/10.1186/s13104-018-3730-x>
- 507 32. Barkin, JL. The Barkin Index of Maternal Functioning. n.d.  
508 <https://barkinindexmaternalfunctioning.com>