

1     **A configural model of expert judgement as a preliminary epidemiological**  
2             **study of injury problems: An application to drowning.**

3

4

5                     Damian Morgan<sup>1\*</sup>, Joan Ozanne-Smith<sup>2</sup>

6

7

8

9     <sup>1</sup> Federation Business School, Federation University Australia, Churchill, Victoria, Australia.

10

11     <sup>2</sup> Department of Forensic Medicine, Monash University, Melbourne, Victoria, Australia.

12

13     \* Corresponding author

14

15     Email: [damian.morgan@federation.edu.au](mailto:damian.morgan@federation.edu.au)

16

17

18

19     Conceived and designed the study: DM JO-S. Collected and analysed data: DM. Wrote the

20     paper: DM JO-S.

21

22

23

24

25

26

27

28

29 **A configural model of expert judgement as a preliminary epidemiological study of**  
30 **injury problems: An application to drowning.**

31

32 **Abstract**

33 Robust epidemiological studies identifying determinants of negative health outcomes require  
34 significant research effort. Expert judgement is proposed as an efficient alternative or  
35 preliminary research design for risk factor identification associated with unintentional injury.  
36 This proposition was tested in a multi-factorial balanced experimental design using specialist  
37 judges (N=18), lifeguards and surfers, to assess the risk contribution to drowning for  
38 swimming ability, surf bathing experience, and wave height. All factors provided unique  
39 contributions to drowning risk ( $p<.001$ ). An interaction ( $p=.02$ ) indicated that occasional surf  
40 bathers face a proportionally increased risk of drowning at increased wave heights relative to  
41 experienced surf bathers. Although findings were limited by strict criteria, and no gold  
42 standard comparison data were available, the study provides new evidence on causal risk  
43 factors for a drowning scenario. Countermeasures based on these factors are proposed.  
44 Further application of the method may assist in developing new interventions to reduce  
45 unintentional injury.

## 47 **Introduction**

48 Analytic epidemiological studies test for the association of determinants with a negative  
49 health outcome to support a theory of causality. Identified causal risk factors may then be  
50 modified to improve health outcomes. Epidemiological research designs provide robust  
51 evidence through observing candidate risk factors in the natural course of events. Means of  
52 potential risk factor identification and specification include anecdotal evidence, case reports  
53 or cross-sectional surveys.

54

55 A significant challenge for observational epidemiological studies concerns the control of  
56 confounding factors [1]. Elimination of competing explanations for study findings often  
57 requires substantial study sizes. Accurate measurement of exposure to risk factors is complex  
58 as is distinguishing factor causation from statistical association. Without attention to these  
59 details, epidemiologic studies provide little substantial knowledge gain.

60

61 Uncertainty about the derived health benefits leads to difficulty in justifying study costs and  
62 may partially explain the lack of epidemiological studies for many significant injury  
63 problems. To provide evidence on potential health gains, cost-effective preliminary studies  
64 with acceptable internal validity and generalisability may guide later epidemiological  
65 research. Such studies should replicate more rigorous epidemiological designs with respect to  
66 risk factor identification and assessment. The study reported here tests and evaluates a  
67 proposed method based on expert opinion for potential causal injury risk factor identification  
68 and risk quantification as applied to unintentional drowning in an Australian surf bather  
69 population, where surf bather drowning is a relative rare event relative to bather numbers [2].

## 70 **Drowning as a global health problem**

71 Globally, drowning accounts for more than 370,000 deaths each year [3]. The causes of  
72 drowning are complex and risk factors vary by geographic location, physical environmental  
73 features, activity, water entry mechanism, weather and water conditions, supervision, and  
74 personal characteristics. Therefore, studies of causal risk factors should be restricted to  
75 clearly defined circumstances.

76

77 In most drowning scenarios, including surf bather drowning, scant evidence exists on causal  
78 risk factors. Surf bathing at wave-dominated beaches attracts local residents and is commonly  
79 depicted in tourist brochures luring visitors to warm and exotic beach locations [4]. Despite  
80 dedicated beach patrols and lifeguards supervising bathers, Australia's annual coastal  
81 drowning rate remains 0.14 swimmers and waders per 100,000 resident population [5].  
82 Several causal drowning risk factors have been proposed for surf bathers, including rip  
83 currents, alcohol, tourists, and onset of medical conditions [6]. Yet no rigorous research  
84 studies confirm that these or other candidate causal risk factors place bathers at relatively  
85 higher drowning risk.

86

87 This study aimed to test the capacity of experts to identify and assess the roles played by  
88 putative causal risk factors in surf bather drowning.

## 89 **Judgement and risk assessment**

90 Expert judgement provides a recognised method for gathering evidence where traditional  
91 scientific methods are impractical [7-8], such as assigning risk probabilities to assess certain  
92 environmental hazards [9-10]. Such risk assessments may be biased by judge overconfidence,  
93 inaccuracy, or insufficient or irrelevant judge expertise [11-12]. Given these and other

94 potential limitations, experiments based on subjective judgements require control of  
95 recognised potential bias and careful selection of judges, thus limiting generalisability of the  
96 findings.

97

98 Early judgement studies sought to model the decision-making process and assess the  
99 applicability of outcomes to the *true state* [13-15]. Linear modelling has matched the process  
100 used by judges where variables are assigned weights, with the sum used to determine the  
101 outcome likelihood. In reality, judges may follow a configural rather than a linear process by  
102 assigning values to predictor variables based on weights of other predictor variables [14].  
103 Related statistical techniques such as analysis-of-variance (ANOVA) can account for  
104 configularity within judgements (predictor variable interactions).

## 105 **Method**

106 Fig 1 presents the study design overview. Based on a repeated-measures multi-factorial  
107 experimental design, two separate groups of surf bathing *specialists*—lifeguards and  
108 surfers—were recruited to judge the contribution of putative causal factors to surf bather  
109 drowning risk.

110 **Fig 1. Study design overview.**

## 111 **Independent variables and factor levels**

112 A preliminary study eliciting water safety expert knowledge, using a nominal group  
113 technique, reported swimming ability in surf conditions, awareness of surf hazards, and  
114 prevailing surf conditions as the top ranked factors affecting the probability of surf bather  
115 drowning [16]. From this, three independent (predictor) variables (IVs) were specified as  
116 putative drowning risk factors: *swimming ability*, *surf bathing experience*, and *wave height*.

117

118 IVs were set at three fixed levels representing ordinal scales, restricting inferences to the  
119 specified IV levels [17]. Factor levels were distinguished by a mix of qualitative descriptions  
120 and quantitative measures (S1 Table). Levels were ordered from (presumed) lowest to highest  
121 drowning risk contribution, assuming other factors are absent (e.g., alcohol), remain equal or  
122 constant (e.g., tide level or health status). For each factor, the median risk level was anchored  
123 at averages found for Australian beaches or surf bather populations [18-19].

## 124 **Dependent (criterion) variable**

125 A scale measurement was required for surf bather drowning risk. Piloting revealed judge  
126 preference for the term ‘getting into difficulty’ as a proxy scale measure of drowning risk.  
127 This scale appraised the likelihood of bathers reaching their limit to cope with surf conditions  
128 based on their swimming ability and surf bathing experience and logically, this situation is a  
129 precursor to drowning. This scale was used as the proxy drowning risk measure for the study.

130

131 The dependent (criterion) variable (DV) used an 11-point scale to record the perceived  
132 chance of getting into difficulty while bathing—0% to 100% [20]. Descriptive terms below  
133 the scale qualified associated percentage ranges for getting into difficulty: No; low; moderate;  
134 high chance; and certain (S1 Table).

## 135 **Hypotheses**

136 Three hypotheses (Hs) were specified:

137

138 H1: *The IVs—swimming ability, surf bathing experience, and wave height—will each*  
139 *produce an effect on the DV—chance of getting into difficulty in the water.*

140

141 Each of the three scale items was expected to be associated with varying levels of surf bather  
142 drowning risk in a systematic *risk* order providing the rationale for H2.

143

144 H2: *The order of levels within each IV is associated systematically with surf bather drowning*  
145 *risk.*

146

147 Null-hypotheses tested for differences between IVs levels and expected direction.

148

149 Intuitively, and consistent with a configural judgement approach, the drowning risk for levels  
150 of one IV would be expected to be influenced by the other two IV levels. Therefore, H3  
151 anticipated that an effect produced on the DV by one level of an IV is dependent on levels of  
152 other IVs [21].

153

154 H3: *Three first order interactions and one second order interaction among the IVs will*  
155 *produce an effect on the DV.*

156

157 H3 generated four null-hypotheses. The direction of interaction effects between IV levels  
158 were investigated *a priori* but not as specified hypotheses.

## 159 **Instrument and design**

160 Ethical approval was granted by the Monash University Standing Committee on Ethics in  
161 Research Involving Humans. The experiment was administered using a self-completed  
162 questionnaire. Personal data comprised surf bathing experience and currency, surf-activity  
163 proficiency, lifesaver/lifeguard and rescue experience and demographic details. To reduce  
164 potential influence of other possible risk factors on judgements, an instruction page outlined

165 the general scenario for drowning risk exposure including bathing at the outer wave breaking  
166 zone (S2 Table). Following this, 27 vignettes, on separate pages, provided a combination of  
167 IV levels and the DV drowning risk scale. Two sets of questionnaires were produced; P1 for  
168 time-period 1 and P2 for time-period 2. Respondents were instructed not to refer back to their  
169 previous ratings when rating new scenarios.

## 170 **Vignettes and ordering procedure**

171 The three IVs at three levels resulted in 27 unique combinations (cells) for rating the DV.  
172 Each cell was presented as a three paragraph vignette personalised with a gendered name (S2  
173 Table).

174

175 The order of IVs can affect the DV score due to participant practice, fatigue or becoming  
176 *wise* to the experiment [21]. To counterbalance carry-over (order) effects, a Latin square type  
177 arrangement was used to distribute carry-over effects systematically across cells [17, 22]. It  
178 was anticipated that statistical analysis would account for carry-over effects within error  
179 terms. The sequence of IVs and subject gender within each vignette was also systematically  
180 varied.

181

182 A repeat square for P2 provided supplemental data to assess the reliability of responses  
183 between P1 and P2. Effectively, the use of the repeat square provided two cell ratings per  
184 judge under different order conditions [23].

## 185 **Population and sampling procedure**

186 Expertise encompasses skills and knowledge and experts may be identified for specific areas  
187 from characteristics including capabilities, achievements, qualifications, peer recognition,  
188 specialisation or years of performance [24-25]. Two populations were considered expert in



189 surf bathing activities, professional lifeguards and proficient surfers, allowing comparison of  
190 judgements across specialist populations. Through bather supervision, lifeguards have direct  
191 experience of bathers getting into difficulty and hold recognised proficiency in surf bathing  
192 activities. Surfer expertise encompasses the necessary skills and experience to negotiate  
193 typical and atypical surf conditions.

## 194 **Sample size and statistical power**

195 Stevens [26] provides required sample sizes for single group repeated-measures ANOVA. To  
196 obtain 80 percent power, assuming an average correlation of DV measures being 0.5, three  
197 treatments (IVs), alpha level of 0.05, and large effect size, 8-14 repeated-measures are  
198 required. The target sample size was nine lifeguard and nine surfer judges. Statistical power  
199 may be increased through pooling of results where no statistical differences are found  
200 between the specialist groups or time periods.

## 201 **Selection and study participation**

202 Participants were selected using a convenience sampling procedure following a snowball-like  
203 process. Some participants were known to each other. All questionnaires were completed in  
204 DM's presence. Following instruction, 18 participants completed the first set of 27 vignettes  
205 (each in a unique order) followed by a 30 minute break. Demographic information and the  
206 second set (repeat square) were then completed.

## 207 **Data analysis**

208 Introduced bias was firstly assessed for vignette gender, time-period effect, and specialist  
209 type. Following this, factorial repeated-measures three-way ANOVA established simple main  
210 effects and error terms for the IVs (H1). Statistical differences for IV levels and direction

211 were then assessed (H2). Planned polynomial contrasts between IVs specified interactions  
212 (H3).

213

214 The ANOVA results follow the order of hypotheses. Where an interaction between IVs was  
215 found, simple pairwise comparisons of means were examined *post-hoc* using the middle IV  
216 level (median risk level) as the reference group [27].

217

218 Preliminary data analyses revealed a wording error for Vignette 7—one IV factor level being  
219 incorrect. This resulted in estimated V7 DV scores for each judge being made by  
220 interpolation [28]. This took into account mean score patterns for corresponding vignette  
221 levels and applied differences to individual V7 scores [29]. The procedure maintained  
222 existing order effects and parallel risk assessments particular to each judge. Specifically, the  
223 score for V7 (with same procedure followed for the repeat square) was calculated as equal to:  
224 original score (V7) less mean difference between V9 and V8 less mean difference between  
225 V17 and V16. This resulted in 3 of 18 judges' scores for P1 being negative (-1.2, -0.2, -0.2)  
226 and 1 judge for P2 (-0.5). These four negative scores were converted to zero. The face  
227 validity for the derived V7 mean score was confirmed by comparison with the *closest* cell  
228 levels. Remaining reporting treats the derived estimates for V7 as the true scores.

229

230 Data were entered on the spreadsheet and analysed using statistical software [30] with alpha  
231 level  $p < 0.05$ . The DV results were entered as scores (0-10) corresponding to percentage  
232 indicated. Normality of distributions was assessed for each vignette visually and by reference  
233 to significance tests for skewness and kurtosis z-scores ( $p < 0.05$ ). Due to small sample sizes  
234 and potential non-normal distributions, non-parametric tests (exact significance) were used  
235 for preliminary subgroup comparisons.

236

237 Differences between specialist groups on demographic and beach behaviour were determined  
238 by Mann-Whitney [ $U$ ] tests or chi-square [ $\chi^2$ ] with corresponding effect size calculated  
239 manually for significant results. To test for vignette name gender effects, P1 and P2 data were  
240 combined for each of the 27 vignettes (each being rated 36 times, twice by a single judge)  
241 with differences assessed by Mann-Whitney ( $U$ ) tests.

242

243 Reliability of judges' scores over P1 and P2 was assessed by the Wilcoxon signed-rank test.  
244 Here, significant results at  $p < 0.05$  would be expected as the cumulative (family-wise) Type 1  
245 error rate across 27 tests increased the likelihood of false positives. Bonferroni correction set  
246 a significant alpha level at  $p < 0.002$  [31]. DV ratings for each vignette were grouped for P1  
247 and P2 to test for differences between specialist groups using Mann-Whitney ( $U$ ) tests.

248

249 Following preliminary assessment, a repeated-measures three-way ANOVA was run on SPSS  
250 using the general linear program. The assumption of sphericity was assessed by Mauchly's  
251 test. Greenhouse-Geisser estimate, correcting for degrees of freedom, was used where the  
252 assumption of sphericity was not met. *A priori* polynomial contrasts were specified for each  
253 factor to test for presumed IV factor order (in linear or quadratic form) with results reported  
254 where significant. Tests of differences between estimated marginal means for IV factor levels  
255 (i.e., the unweighted mean that controls for potential confounding from other IVs), three first-  
256 order interactions, and one second-order interaction, applied the Bonferroni correction.  
257 Partial eta squared (partial  $\eta^2$ ), which explains the proportion of variation unique to a variable  
258 not explained by other variables, was used to estimate effect sizes [27] with 95% CI  
259 calculated from SPSS syntax files from Smithson cited in [21].

260

261 Eta squared ( $\eta^2$ ) was calculated manually. This measure can be interpreted in a similar way to  
262  $R^2$ , being an additive portion of the total variance in the DV explained by the IVs and  
263 interactions, provided the design is balanced by an equal group size for each cell [21-22].  
264 These tests for effect size were based on the sample results without correction for population  
265 estimates [27]. Figures were prepared manually using the Excel program [32].

## 266 **Results**

### 267 **Specialist profiles**

268 The specialist groups (lifeguards and surfers) had similar demographics and beach experience  
269 confirmed by non-significant differences on statistical tests (not reported). Surfers had higher  
270 frequency of beach visits in the previous 12 months ( $U=18.5$ ,  $p=0.05$ ). All judges had  
271 extensive experience in surf bathing (10 to 30 years). Most had high participation rates in the  
272 last 12 months and experience in 3 m waves. All judges except one rated themselves as  
273 proficient or expert in surf bathing.

274

275 All lifeguards held surf-related and first aid qualifications and had completed rescues, six  
276 having performed cardiopulmonary resuscitation. Almost half the surfer specialists held  
277 swimming-related qualifications, first aid certification or experience in performing rescues. A  
278 statistically significant difference was found for the average number of rescues performed by  
279 surfers and lifeguards (respective means 2.3 and 270.6,  $U<0.01$ ,  $p<0.01$ ,  $r= -1.2$ ).

### 280 **Vignette gender**

281 The DV mean rating (chance of getting into difficulty in the water) was higher for female  
282 vignette subjects compared to males for 17 vignette scenarios (63%); lower for 10 (37%).

283 Mann-Whitney ( $U$ ) tests were not significant for any vignette gender differences ( $p < 0.05$ ) so  
284 this variable was not treated as a factor in further analysis.

## 285 **Reliability of vignette DV ratings between P1 and P2**

286 Vignette DV mean ratings between P1 and P2 ( $N=36$ ) were 13 (48%) higher cell means for  
287 P1, 12 (44%) higher cell means for P2, and 2 (7%) identical means. As Wilcoxon sign-rank  
288 test identified no significant differences ( $p < 0.002$ ) P1 and P2 ratings were considered  
289 statistically to be from the same populations, providing justification for pooling cell ratings.

## 290 **Comparison of specialist groups DV ratings**

291 Table 1 presents judges' mean vignette ratings for P1, P2 and overall. Although vignette  
292 ratings varied (mean 4.0-6.9), when grouped the pattern for surfers and lifeguards were  
293 similar. Overall mean differences between specialist group ratings were not significant for  
294 P1, P2, or combined periods (P1:  $U=35.0$ ,  $p=0.65$ ,  $r=-0.11$ . P2:  $U=27.0$ ,  $p=0.25$ ,  $r=-0.28$ ,  
295 combined:  $U=30.5$ ,  $p=0.40$ ,  $r=-0.21$ ).

296

297

298

299

300

301

302

303

304

305

306

307 **Table 1.** Specialist judges mean ratings of the DV for 27 vignette scenarios.

Judge	Specialist	P1		P2		Combined periods		Range
		Mean	SD	Mean	SD	Mean	SD	
1	Surfer	6.8	2.4	7.0	2.8	6.9	2.6	2-10
2	Surfer	5.4	3.2	5.7	3.2	5.6	3.2	0-10
3	Surfer	6.7	2.9	6.1	3.1	6.4	3.0	1-10
4	Surfer	6.5	3.4	6.4	3.7	6.4	3.5	0-10
5	Surfer	6.1	3.1	6.0	3.4	6.1	3.2	1-10
6	Surfer	5.2	2.5	5.4	3.0	5.3	2.7	1-10
7	Surfer	4.3	3.0	4.1	2.5	4.2	2.8	0-10
8	Surfer	5.2	2.5	5.6	2.8	5.4	2.6	1-10
9	Surfer	3.9	3.3	5.5	3.3	4.7	3.4	0-10
	<i>Total surfers</i>	5.6	3.0	5.7	3.2	5.7	3.1	0-10
10	Lifeguard	6.9	2.7	7.0	2.8	6.9	2.7	1-10
11	Lifeguard	4.7	2.6	4.4	2.3	4.5	2.4	0-10
12	Lifeguard	6.7	3.0	6.5	3.3	6.6	3.1	1-10
13	Lifeguard	3.9	1.8	4.1	1.7	4.0	1.7	1-7
14	Lifeguard	4.0	2.8	4.7	2.6	4.3	2.7	0-10
15	Lifeguard	4.8	2.2	4.5	2.1	4.7	2.1	1-8
16	Lifeguard	4.4	2.7	4.4	2.8	4.4	2.7	1-9
17	Lifeguard	5.8	2.6	5.2	2.2	5.5	2.4	0-10
18	Lifeguard	6.5	2.4	5.8	2.5	6.1	2.4	1-10
	<i>Total lifeguards</i>	5.3	2.7	5.2	2.7	5.2	2.7	0-10
	<i>Overall total</i>	5.4	2.9	5.5	2.9	5.4	2.9	0-10

308 *Note:* DV: The chance of getting into difficulty while bathing (scale 0—no chance to 10—certain)

309

310

311 DV rating patterns for individual vignette cells by specialist group provide further insight

312 (Table 2). Table 2 shows broad similarity between surfer and lifeguard DV patterns across

313 vignette cells. Upper and lower DV ratings for individual cells ranged substantially, partly

314 explained by order effects given a small drop in the overall standard deviation from P1 to P2.

315

316

317 **Table 2.** Specialist judges mean ratings of DV for vignette cells by group.

Vignette 1-13	IV levels	Specialist	P1		P2		Combined periods		Range
			Mean	SD	Mean	SD	Mean	SD	
1	Swim.cap 1	Surfer	1.0	0.9	0.7	0.7	0.8	0.8	0-3
	Surf.exp 1	Lifeguard	1.3	0.7	1.4	1.0	1.4	0.9	0-4
	Waves 1	<i>Overall</i>	<i>1.1</i>	<i>0.8</i>	<i>1.1</i>	<i>1.0</i>	<i>1.1</i>	<i>0.9</i>	<i>0-4</i>
2	Swim.cap 1	Surfer	3.7	1.9	3.7	1.9	3.7	1.9	1-7
	Surf.exp 1	Lifeguard	4.2	2.3	4.1	2.1	4.1	2.1	1-8
	Waves 2	<i>Overall</i>	<i>3.9</i>	<i>2.1</i>	<i>3.9</i>	<i>2.0</i>	<i>3.9</i>	<i>2.0</i>	<i>1-8</i>
3	Swim.cap 1	Surfer	5.8	2.9	6.3	2.7	6.1	2.8	1-10
	Surf.exp 1	Lifeguard	5.4	2.1	5.3	2.0	5.2	2.0	3-9
	Waves 3	<i>Overall</i>	<i>5.6</i>	<i>2.5</i>	<i>5.8</i>	<i>2.4</i>	<i>5.7</i>	<i>2.4</i>	<i>1-10</i>
4	Swim.cap 1	Surfer	1.6	1.1	0.9	0.7	1.3	0.9	0-4
	Surf.exp 2	Lifeguard	1.8	1.2	1.9	1.1	1.9	1.1	1-4
	Waves 1	<i>Overall</i>	<i>1.7</i>	<i>1.1</i>	<i>1.4</i>	<i>1.0</i>	<i>1.6</i>	<i>1.1</i>	<i>0-4</i>
5	Swim.cap 1	Surfer	4.9	1.8	5.8	2.1	5.4	2.0	1-8
	Surf.exp 2	Lifeguard	5.0	2.4	5.5	1.9	5.3	2.1	2-9
	Waves 2	<i>Overall</i>	<i>4.9</i>	<i>2.1</i>	<i>5.7</i>	<i>1.9</i>	<i>5.3</i>	<i>2.0</i>	<i>1-9</i>
6	Swim.cap 1	Surfer	6.4	1.8	7.8	2.4	7.1	2.2	2-10
	Surf.exp 2	Lifeguard	5.6	2.3	6.3	1.3	5.9	1.9	2-9
	Waves 3	<i>Overall</i>	<i>6.0</i>	<i>2.1</i>	<i>7.1</i>	<i>2.0</i>	<i>6.5</i>	<i>2.1</i>	<i>2-10</i>
7	Swim.cap 1	Surfer	2.9	1.5	3.2	0.7	3.1	1.1	0-5
	Surf.exp 3	Lifeguard	2.2	1.8	1.4	1.7	1.8	1.8	0-5
	Waves 1	<i>Overall</i>	<i>2.6</i>	<i>1.6</i>	<i>2.3</i>	<i>1.6</i>	<i>2.4</i>	<i>1.6</i>	<i>0-5</i>
8	Swim.cap 1	Surfer	5.7	1.5	6.9	1.5	6.3	1.6	3-9
	Surf.exp 3	Lifeguard	5.8	1.9	5.9	1.5	5.9	1.7	3-9
	Waves 2	<i>Overall</i>	<i>5.6</i>	<i>1.7</i>	<i>6.4</i>	<i>1.5</i>	<i>6.1</i>	<i>1.6</i>	<i>3-9</i>
9	Swim.cap 1	Surfer	7.4	1.5	8.9	1.1	8.2	1.5	5-10
	Surf.exp 3	Lifeguard	7.6	1.8	7.6	1.2	7.6	1.5	4-10
	Waves 3	<i>Overall</i>	<i>7.5</i>	<i>1.6</i>	<i>8.2</i>	<i>1.3</i>	<i>7.9</i>	<i>1.5</i>	<i>4-10</i>
10	Swim.cap 2	Surfer	1.1	0.8	1.2	0.4	1.1	0.6	0-2
	Surf.exp 1	Lifeguard	1.9	0.7	2.0	1.3	2.0	1.0	1-5
	Waves 1	<i>Overall</i>	<i>1.5</i>	<i>0.8</i>	<i>1.6</i>	<i>1.0</i>	<i>1.6</i>	<i>0.9</i>	<i>0-5</i>
11	Swim.cap 2	Surfer	4.7	2.3	4.9	1.9	4.8	2.1	0-8
	Surf.exp 1	Lifeguard	4.6	2.2	4.7	2.0	4.7	2.1	1-8
	Waves 2	<i>Overall</i>	<i>4.6</i>	<i>2.2</i>	<i>4.8</i>	<i>1.9</i>	<i>4.7</i>	<i>2.0</i>	<i>0-8</i>
12	Swim.cap 2	Surfer	6.8	2.0	7.1	1.8	6.9	1.9	4-10
	Surf.exp 1	Lifeguard	5.9	2.8	6.2	2.1	6.1	2.4	1-9
	Waves 3	<i>Overall</i>	<i>6.4</i>	<i>2.4</i>	<i>6.6</i>	<i>2.0</i>	<i>6.5</i>	<i>2.2</i>	<i>1-10</i>
13	Swim.cap 2	Surfer	1.9	1.2	1.9	0.9	1.9	1.0	0-4
	Surf.exp 2	Lifeguard	2.4	1.2	2.2	0.9	2.3	1.0	1-5
	Waves 1	<i>Overall</i>	<i>2.2</i>	<i>1.2</i>	<i>2.1</i>	<i>0.9</i>	<i>2.1</i>	<i>1.0</i>	<i>0-5</i>

318 *Note:* DV: The chance of getting into difficulty while bathing (scale 0—no chance to 10—certain)  
319 IV Swim.cap—swimming ability (1=lowest risk, 2=middle risk, 3=highest risk)  
320 IV Surf.exp—surf bathing experience (1=lowest risk, 2=middle risk, 3=highest risk)  
321 IV Waves—wave height (1=lowest risk, 2=middle risk, 3=highest risk)  
322 Vignette 7 scores derived via interpolation following researcher error.

323 **Table 2.** (*continued*) Specialist judges mean ratings of DV for vignette cells by group.

Vignette	IV levels	Specialist	P1		P2		Combined periods		Range
			Mean	SD	Mean	SD	Mean	SD	
14	Swim.cap 2	Surfer	5.9	1.5	6.1	1.5	6.0	1.5	3-9
	Surf.exp 2	Lifeguard	5.4	1.4	5.2	1.8	5.3	1.6	3-8
	Waves 2	<i>Overall</i>	<i>5.7</i>	<i>1.4</i>	<i>5.7</i>	<i>1.7</i>	<i>5.7</i>	<i>1.5</i>	<i>3-9</i>
15	Swim.cap 2	Surfer	8.0	1.4	8.2	1.5	8.1	1.4	5-10
	Surf.exp 2	Lifeguard	6.6	1.9	7.0	1.6	6.8	1.8	5-10
	Waves 3	<i>Overall</i>	<i>7.3</i>	<i>1.8</i>	<i>7.6</i>	<i>1.6</i>	<i>7.4</i>	<i>1.7</i>	<i>5-10</i>
16	Swim.cap 2	Surfer	3.2	2.1	2.3	0.8	2.8	1.6	1-8
	Surf.exp 3	Lifeguard	3.0	1.1	2.7	1.1	2.8	1.1	1-5
	Waves 1	<i>Overall</i>	<i>3.1</i>	<i>1.6</i>	<i>2.5</i>	<i>0.9</i>	<i>2.8</i>	<i>1.3</i>	<i>1-8</i>
17	Swim.cap 2	Surfer	6.7	1.9	7.3	1.5	7.0	1.7	4-9
	Surf.exp 3	Lifeguard	6.5	1.4	6.0	1.9	6.3	1.6	3-8
	Waves 2	<i>Overall</i>	<i>6.6</i>	<i>1.6</i>	<i>6.7</i>	<i>1.8</i>	<i>6.6</i>	<i>1.7</i>	<i>3-9</i>
18	Swim.cap 2	Surfer	8.3	1.5	8.7	1.3	8.5	1.4	6-10
	Surf.exp 3	Lifeguard	8.1	1.2	7.7	1.7	7.9	1.4	5-10
	Waves 3	<i>Overall</i>	<i>8.2</i>	<i>1.3</i>	<i>8.2</i>	<i>1.5</i>	<i>8.2</i>	<i>1.4</i>	<i>5-10</i>
19	Swim.cap 3	Surfer	3.2	2.0	3.4	1.7	3.3	1.8	1-7
	Surf.exp 1	Lifeguard	3.4	1.8	2.9	0.8	3.2	1.4	1-6
	Waves 1	<i>Overall</i>	<i>3.3</i>	<i>1.9</i>	<i>3.2</i>	<i>1.3</i>	<i>3.2</i>	<i>1.6</i>	<i>1-7</i>
20	Swim.cap 3	Surfer	7.2	1.5	7.1	1.5	7.2	1.5	4-9
	Surf.exp 1	Lifeguard	6.2	1.5	5.9	1.5	6.1	1.5	4-9
	Waves 2	<i>Overall</i>	<i>6.7</i>	<i>1.6</i>	<i>6.5</i>	<i>1.6</i>	<i>6.6</i>	<i>1.5</i>	<i>4-9</i>
21	Swim.cap 3	Surfer	8.3	2.2	9.0	1.4	8.7	1.8	4-10
	Surf.exp 1	Lifeguard	7.9	1.7	7.7	1.6	7.8	1.6	6-10
	Waves 3	<i>Overall</i>	<i>8.1</i>	<i>1.9</i>	<i>8.3</i>	<i>1.6</i>	<i>8.2</i>	<i>1.7</i>	<i>4-10</i>
22	Swim.cap 3	Surfer	3.8	1.3	3.3	1.3	3.6	1.3	2-7
	Surf.exp 2	Lifeguard	4.1	1.8	3.5	1.2	3.8	1.5	2-8
	Waves 1	<i>Overall</i>	<i>3.9</i>	<i>1.5</i>	<i>3.4</i>	<i>1.2</i>	<i>3.7</i>	<i>1.4</i>	<i>2-8</i>
23	Swim.cap 3	Surfer	8.6	0.9	7.9	1.6	8.2	1.3	5-10
	Surf.exp 2	Lifeguard	7.3	1.2	6.9	1.6	7.1	1.4	5-10
	Waves 2	<i>Overall</i>	<i>7.9</i>	<i>1.2</i>	<i>7.4</i>	<i>1.6</i>	<i>7.7</i>	<i>1.4</i>	<i>5-10</i>
24	Swim.cap 3	Surfer	9.4	0.9	9.4	0.9	9.4	0.8	7.5-10
	Surf.exp 2	Lifeguard	8.8	1.2	8.4	1.5	8.6	1.3	5.5-10
	Waves 3	<i>Overall</i>	<i>9.1</i>	<i>1.1</i>	<i>8.9</i>	<i>1.3</i>	<i>9.0</i>	<i>1.2</i>	<i>5.5-10</i>
25	Swim.cap 3	Surfer	5.1	2.4	4.1	1.6	4.6	2.1	2-9
	Surf.exp 3	Lifeguard	4.7	1.6	4.3	1.1	4.5	1.4	2-8
	Waves 1	<i>Overall</i>	<i>4.9</i>	<i>2.0</i>	<i>4.2</i>	<i>1.4</i>	<i>4.6</i>	<i>1.7</i>	<i>2-9</i>
26	Swim.cap 3	Surfer	9.0	1.3	9.1	0.3	9.1	0.9	7-10
	Surf.exp 3	Lifeguard	8.2	1.6	7.9	1.4	8.1	1.5	5-10
	Waves 2	<i>Overall</i>	<i>8.6</i>	<i>1.5</i>	<i>8.5</i>	<i>1.2</i>	<i>8.6</i>	<i>1.3</i>	<i>5-10</i>
27	Swim.cap 3	Surfer	10.0	0.0	10.0	0.0	10.0	0.0	10-10
	Surf.exp 3	Lifeguard	8.9	1.1	9.0	1.1	8.9	1.1	7-10
	Waves 3	<i>Overall</i>	<i>9.4</i>	<i>0.9</i>	<i>9.5</i>	<i>0.9</i>	<i>9.5</i>	<i>0.9</i>	<i>7-10</i>

324 *Note:* DV: The chance of getting into difficulty while bathing (scale 0—no chance to 10—certain)  
325 IV Swim.cap—swimming ability (1=lowest risk, 2=middle risk, 3=highest risk)  
326 IV Surf.exp—surf bathing experience (1=lowest risk, 2=middle risk, 3=highest risk)  
327 IV Waves—wave height (1=lowest risk, 2=middle risk, 3=highest risk)



328 Applying the Bonferroni correction ( $p < 0.002$ ) left vignette 27 as the only statistically  
329 significant difference between specialists groups. Thus surfers' and lifeguards' ratings of the  
330 DV by vignette IV order levels were considered to be from the same population of specialists.  
331 Fig 2 shows the overall pattern of combined means for each of the 27 vignettes.

332

333 **Fig 2. Estimated marginal mean drowning risk scores on DV for swimming ability by**  
334 **surf bathing experience and wave height.**

335

336 The repeated-measures ANOVA procedure is robust to violations in the normality  
337 assumption of DV distributions [26]. Nevertheless, each vignette distribution for combined  
338 periods and judges was assessed for normality. Visual appearance approximated normal  
339 distributions. No distribution was significantly skewed ( $p < 0.05$ ). Four (14.8%) vignette cell  
340 distributions were significant for kurtosis ( $p < 0.05$ ) due to a high peak score (many judges  
341 chose the same rating score). Based on these results, the DV data were considered suitable for  
342 further analysis without transformation.

### 343 **Repeated-measures factorial ANOVA**

344 A repeated-measures three-way ANOVA determined significant effects and polynomial  
345 contrasts ( $p < 0.05$ ) between the chance of getting into difficulty in the water (DV) and IVs  
346 swimming ability, surf bathing experience, and wave height, and IV interactions. Table 3 lists  
347 marginal means and standard error scores on the DV for each level of the three IVs. The  
348 overall hypothesised pattern of drowning risk posed by IV factor levels was reflected in the  
349 expert ratings.

## 350 Hypotheses tests

351 The model resulted in simple main effects for the three IVs on the DV; swimming ability,  
352  $F(1.21, 20.62) = 77.87, p < 0.001$ , partial  $\eta^2 = 0.82$  (95% CI: 0.62 to 0.88), surf bathing  
353 experience,  $F(2, 34) = 99.27, p < 0.001$ , partial  $\eta^2 = 0.85$  (95% CI: 0.74 to 0.90), and wave  
354 height,  $F(1.30, 22.05) = 227.95, p < 0.001$ , partial  $\eta^2 = 0.93$  (95% CI: 0.85 to 0.95). All levels  
355 within each IV differed following pairwise comparisons with Bonferroni correction ( $p < .001$ ;  
356 Table 3). Polynomial contrasts revealed a significant linear trend for the three IVs; swimming  
357 ability,  $F(1,17) = 88.51, p < 0.001$ , partial  $\eta^2 = 0.84$ , surf bathing experience,  $F(1,17) = 135.01$ ,  
358  $p < 0.001$ , partial  $\eta^2 = 0.89$ , and wave height,  $F(1,17) = 5277.04, p < 0.001$ , partial  $\eta^2 = 0.94$  with  
359 significant quadratic trends for swimming ability,  $F(1,17) = 32.07, p < 0.001$ , partial  $\eta^2 = 0.65$   
360 and wave height,  $F(1,17) = 52.43, p < 0.001$ , partial  $\eta^2 = 0.76$ . A significant effect was found  
361 for the interaction of surf bathing experience and wave high,  $F(4, 68) = 3.03, p = 0.02$ , partial  
362  $\eta^2 = 0.15$  (95% CI: 0.00 to 0.27). Fig 3 shows the marginal mean scores. Polynomial  
363 contrasts found a significant linear interaction within the quadratic pattern for waves,  $F(1, 17)$   
364  $= 8.15, p = 0.01$ , partial  $\eta^2 = 0.32$ . A *Post-hoc* contrast comparing experienced beach  
365 swimmers to occasional beach swimmers in waves 0.5 and 2.0 m showed a significant  
366 interaction,  $F(1, 35) = 17.85, p < 0.001$ , partial  $\eta^2 = 0.34$  (95% CI: 0.10 to 0.52).

367

368 **Fig 3. Estimated marginal mean drowning risk scores for surf bathing experience by**  
369 **wave height.**

370

371

372

373

374

375 **Table 3. Estimated marginal means and standard errors for three drowning risk levels**  
 376 **for IVs.**

		DV rating	
		Mean	SE
<i>Swimming ability</i>			
Level 1	Strong swimmer	4.5	0.3
Level 2	Moderately good swimmer	5.1	0.3
Level 3	Weak swimmer	6.8	0.2
<i>Surf bathing experience</i>			
Level 1	Experienced surf bather	4.6	0.3
Level 2	Occasional surf bather	5.4	0.2
Level 3	First time surf bather	6.3	0.2
<i>Wave height</i>			
Level 1	Waves 0.5 m	2.6	0.2
Level 2	Waves 2.0 m	6.1	0.3
Level 3	Waves 3.0 m	7.7	0.3

377 *Note:* DV: The chance of getting into difficulty while bathing (scale 0—no chance to 10—certain).  
 378 Putative lowest drowning risk lowest for level 1 and highest for level 3.  
 379

### 380 **Model estimation**

381  $\eta^2$  as a measure of contributed ANOVA model variance are reported in Table 4. In total, the  
 382 three IVs and interactions explained 75 percent of the variability in the DV. The effects sizes  
 383 are smaller and in different proportions to partial  $\eta^2$  due to the different base used for  
 384 calculations [21].

385

386

387

388

389 **Table 4. Variation in the DV explained by the IVs and interactions.**

IV or interaction term	$\eta^2$
Swimming ability	0.12
Surf bathing experience	0.06
Wave height	0.57
Swimming ability * Surf bathing experience	<0.01
Swimming ability * Wave height	<0.01
Surf bathing experience * Wave height	<0.01
Swimming ability * Surf bathing experience * Wave height	<0.01
Total variation	0.75

390

## 391 **Discussion**

392 Eighteen lifeguards or surfers, meeting study specifications, were considered suitable judges  
393 of surf bather drowning putative risk factors based on their experience in surf activities. Each  
394 judge rated the likelihood of a person requiring rescue in 27 scenarios in a Latin square  
395 arrangement for unique combinations of three levels for swimming ability, surf bathing  
396 experience and wave height, with replication. Due to similarities of ratings, judges were  
397 considered to be from the same population and data for time-periods were pooled. This  
398 increased the study's statistical power by reducing the proportion of error terms from  
399 presumed carryover effects.

400

401 The study found that swimming ability, surf bathing experience and wave height influenced  
402 the risk of surf bather drowning. This risk reduced when: swimming capability increased, surf

403 bathing experience increased, or wave height decreased. The interaction between surf bathing  
404 experience and wave height suggests that drowning risk to novice surf bathers increases  
405 disproportionately at greater wave heights compared to surf bathers with more experience.

406 No other interactions were found in the model.

407

408 The mean ratings for the 27 uniquely ordered vignettes provided a consistent pattern of rated  
409 drowning risk. This pattern *fell* as expected, imparting face validity to the study's method and  
410 sample. DV mean cell ratings ranged from 11 percent chance of drowning for the IV  
411 combination at the lowest level of presumed risk to 95 percent chance at the highest IV risk  
412 combination. This suggests that judges considered *all* surf bathers to carry drowning risk  
413 (under scenario conditions), regardless of their skills, experience, and surf conditions.

414

415 Simple main effects of the three IVs were significant, meaning that alone each contributed  
416 variance to the DV. H1, *the three IVs will each produce an effect on the DV*, is therefore  
417 supported. Each IV accounted for a high proportion of variance within the DV, ignoring that  
418 shared with other IVs. By converting partial  $\eta^2$  results to percentages, these were 82, 85, and  
419 93 percent for swimming ability surf bathing experience, and wave height respectively.

420 Tabachnick and Fidell (p. 188) suggest that repeated-measures “produces a better guess of the  
421 effect size” compared to using a one-way ANOVA design for each factor [27]. For H2,  
422 significant linear trends for IV levels, in the expected drowning risk order, were found for  
423 each IV. The finding validated the specification of each scale as a presumed predictor of surf  
424 bather drowning risk.

425

426 Although a linear pattern of factor level distribution was strongest for each IV (based on  
427 partial  $\eta^2$ ), swimming ability and wave height also formed quadratic patterns within factors.

428 For swimming ability, the increased drowning risk between ratings for a strong swimmer and  
429 moderately good swimmer was less pronounced relative to that between the moderately good  
430 swimmer and the weak swimmer. In contrast, the factor pattern for surf bathing experience  
431 was constant and this IV also showed less total rating variation in the DV compared to the  
432 other two IVs. Wave height showed the greatest difference in estimated marginal means from  
433 lowest to highest risk. The proportionally greater increased rating in drowning risk between  
434 0.5 and 2 m wave height than between 2 and 3 m wave heights supports Short's beach hazard  
435 rating system, where wave height of 0.5 m carries a beach hazard (drowning) rating of 4  
436 (safest), 2.0 m is rated 7, and 3.0 m is rated 9 (least safe) [19].

437

438 For H3, of the three first order interactions between IVs, only that between surf bathing  
439 experience and wave height was significant, with medium effect size. Increasing wave height  
440 from 0.5 to 2.0 m presents greater proportional increase in drowning risk for occasional surf  
441 bathers compared to experienced surf bathers. The second order interaction between the three  
442 IVs was not significant. It was anticipated that drowning protection provided by strong  
443 swimming ability would increase more than proportionally at larger wave heights compared  
444 to less able swimmers. The lack of identified interactions (bar one) suggests that judges  
445 largely rated vignettes in a summative fashion based on an estimated risk contribution at the  
446 specified level for each IV.

447

448 This finding is consistent with previous studies using similar methods [14, 33]. Only one  
449 interaction being identified (assuming interactions exist) may result from judges' failure to  
450 understand the situation correctly or the sensitivity of the research design to identify  
451 configural effects [34]. Alternatively, this result may accurately represent the judgment  
452 process, whether or not this represents the true situation [35].

453

454 Overall, this study provides *some* evidence that surf bathing specialists judge drowning risk  
455 using a configural process. Perhaps the frequent failure of previous studies to identify  
456 significant interactions between variables is explicable by the nature of the judgement task.  
457 Surf bathing experience (human-related factor) and wave height (environmental factor) are  
458 conceptually very different and dynamic variables, yet interactions would be expected. The  
459 identified interaction documents a configural judgement process used by judges to assess  
460 drowning risk from the interplay of environmental and human factors.

## 461 **Limitations**

### 462 **Methodological**

463 Although not reviewed here, statisticians debate the suitability of the long established  
464 repeated-measures factorial ANOVA where other procedures (e.g., MANOVA) may have  
465 less restrictive assumptions or may better model simple interactions [22, 36]. The repeated-  
466 measures ANOVA in this study met required assumptions and provided an appropriate test of  
467 the hypotheses. The decision to group or pool data may be challenged on strict statistical  
468 grounds. For example, a small proportion of comparisons (3.7%) identified differences in  
469 ratings between lifeguards and surfers following Bonferroni corrections. Essentially though,  
470 this limitation was counterbalanced by increased statistical power and dilution of order  
471 effects derived from combining specialist samples.

### 472 **Sample size and selection**

473 Only 18 judges, drawn from a convenience sample, participated. Drowning risk ratings may  
474 be biased by these judges' particular experiences and knowledge. Relevant to the study aim,  
475 the repeated-measures approach has advantage over completely randomised study designs  
476 through requiring fewer participants while having increased power and precision [26]. The

477 consistency between drowning risk ratings from two distinct specialist groups suggests the  
478 study findings may be generalisable all surf bathing experts.

### 479 **Judges' linear interpretation**

480 The reality of vignette scenarios may be questioned. Judges may, for example, have  
481 perceived as contradictory or unrealistic the scenario denoting a subject as a weak swimmer  
482 with extensive surf bathing experience, a form of common method bias [37]. Such a  
483 perception could have encouraged a linear process for rating IVs and so explain the lack of  
484 interaction. At any rate, this study was limited to clearly defined circumstances. Findings  
485 restrict to the influence on drowning risk from three variables specified as fixed factors,  
486 within a general scenario specified (S2 Table), using an untested proxy measure for surf  
487 bather drowning.

### 488 **Unplanned researcher error**

489 One IV level in a single Latin square cell was incorrect. Such data errors or losses are not  
490 unusual in studies with similar research designs where estimating scores is a satisfactory  
491 approach [22] and corrections will not have a disproportional effect on the overall results  
492 [38]. Although the method for estimating missing cell scores for individual judges was crude,  
493 results obtained by correction remained consistent with the mean DV score patterns found for  
494 each of the other 26 vignette ratings.

### 495 **Implications**

496 Based on the study findings, surf bathers are at higher drowning risk, compared to other surf  
497 bathers, where they have inferior swimming ability, less surf bathing experience, or face  
498 larger waves. Although these findings are intuitive, in the absence of robust epidemiological  
499 data, this study provides the best available evidence supporting these conclusions. Regardless  
500 of the study limitations, comparisons between IV mean levels suggest that surf bather



501 drowning results from a complex mix of person and situation variables and provide evidence  
502 on candidate risk factors to guide further investigation of risk and support drowning  
503 prevention strategies.

504

505 International tourists in Australia have a higher rate of surf bather drowning relative to  
506 Australian residents [6] and police reports in coronial records suggest some decedents lacked  
507 experience in surf conditions. Tourist awareness programmes on surf risk and deployment of  
508 lifeguards or surveillance drones to popular tourist areas may mitigate this risk [39-41].

509 Specifically, bathers with little or no surf experience should be aware that strong pool  
510 swimming ability may provide insufficient protection from drowning. Particularly for men,  
511 surf inexperience may translate to overconfidence in one's ability to meet prevailing wave  
512 conditions [18]. Similarly, expansion of surf awareness and safety programmes (e.g. *Nippers*  
513 program for children) on surf beaches during high seasons may contribute to building surf  
514 competency and reducing over-confidence [42-43].

515

516 Additionally, meteorological reports of surf conditions could incorporate indications of risk  
517 level (e.g., not suitable for inexperienced surf swimmers), somewhat similar to current  
518 ultraviolet radiation warnings [44]. In Australia, drowning risk indicators may be integrated  
519 within the Beachsafe website that provides detailed bather-related information for surf  
520 beaches [45]. Technical advances in inflatable lifejackets, originally designed for big wave  
521 surfers, may also offer drowning protection, especially for weak swimmers or inexperienced  
522 surf bathers [46]. Waterproof GPS tracking devices may also potentially aid in bather  
523 surveillance and timely rescue [47]. These and other possible countermeasures require careful  
524 evaluation before their efficacy in drowning prevention is assumed.

525

526 With regard to future epidemiological observations studies, analysis of expert judgments  
527 provides a step towards distinguishing the roles of key variables within this complexity. The  
528 extent to which the findings represent the true surf drowning situation requires comparison to  
529 a gold standard gained only through rigorous epidemiological designs. The availability of  
530 such evidence would be ideal, but meanwhile, the method reported here provides a useful  
531 alternative or preliminary investigation of injury risk factors to reduce drowning.

## 532 **Conclusion**

533 The proof of method reported here offers an important avenue for investigating significant  
534 health problems, including unintentional injury, by providing a *window* into the true situation.  
535 Expert judgement carefully collected and analysed, can be used to document and assess the  
536 roles of risk contributions from putative causal factors that determine health outcomes. This  
537 method provides injury researchers with a rapid low cost tool for data collection comparable  
538 to that obtained through resource intensive epidemiological designs. A method based on  
539 expert judgement of course cannot replace these more robust designs, but may prove a useful  
540 substitute or preliminary method for generating new knowledge to address health problems  
541 and improve outcomes.

542

543

## 544 **Acknowledgments:**

545 The authors express their thanks to the surf specialists who provided their time for  
546 participation in the research study.

## 547 **References**

- 548 1. Susser M. What is a cause and how do we know one? A grammar for pragmatic  
549 epidemiology. *Am J Epidemiol.* 1991;133(7):635-48.
- 550 2. Morgan D, Ozanne-Smith J. Measurement of a drowning incidence rate combining  
551 direct observation of an exposed population with mortality statistics. *International*  
552 *Journal of Injury Control and Safety Promotion.* 2015;22(3):209-14.
- 553 3. World Health Organization. *Global Report on Drowning: Preventing a Leading Killer.*  
554 World Health Organization; 2014.
- 555 4. Morgan D. Counting Beach Visitors: Tools, Methods and Management Applications.  
556 In: Botero CM, Cervantes O, Finkl CW, editors. *Beach Management Tools -*  
557 *Concepts, Methodologies and Case Studies.* Cham: Springer International Publishing;  
558 2018. p. 561-77.
- 559 5. Surf Life Saving Australia. *National Coastal Safety Report 2015.* 2015.
- 560 6. Morgan D, Ozanne-Smith J, Triggs T. Descriptive epidemiology of drowning deaths  
561 in a surf beach swimmer and surfer population. *Inj Prev.* 2008;14(1):62-5. doi:  
562 10.1136/ip.2006.013508.
- 563 7. Dawes RM. The robust beauty of improper linear models in decision making. *Am*  
564 *Psychol.* 1979;34(7):571-82. doi: 10.1037/0003-0066X.34.7.571.
- 565 8. Otway H, Winterfeldt D. Expert judgment in risk analysis and management: process,  
566 context, and pitfalls. *Risk Anal.* 1992;12(1):83-93. doi: 10.1111/j.1539-  
567 6924.1992.tb01310.x.
- 568 9. Fayerweather WE, Collins JJ, Schnatter AR, Hearne FT, Menning RA, Reynefr DP.  
569 Quantifying uncertainty in a risk assessment using human data. *Risk Anal.*  
570 1999;19(6):1077-90. doi: 10.1111/j.1539-6924.1999.tb01129.x.
- 571 10. Morgan MG, Florig HK, DeKay ML, Fischbeck P. Categorizing risks for risk  
572 ranking. *Risk Anal.* 2000;20(1):49-58. doi: 10.1111/0272-4332.00005.
- 573 11. Cox LA. *Risk Analysis: Foundations, Models and Methods.* Norwell: Kluwer  
574 Academic; 2002.
- 575 12. Otway H, Winterfeldt D. Expert judgment in risk analysis and management: process,  
576 context, and pitfalls. *Risk Anal.* 1992;12(1):83-93. doi: 10.1111/j.1539-  
577 6924.1992.tb01310.x.
- 578 13. Hoffman PJ. The paramorphic representation of clinical judgment. *Psychol Bull.*  
579 1960;57(2):116. doi: 10.1037/h0047807.
- 580 14. Hoffman PJ. Cue-consistency and configurality in human judgment. In: Kleinmuntz  
581 B, editor. *Formal Representation of Human Judgement.* New York: Wiley; 1967. p.  
582 53-90.
- 583 15. Hoffman PJ, Slovic P, Rorer LG. An analysis-of-variance model for the assessment of  
584 configural cue utilization in clinical judgment. *Psychol Bull.* 1968;69(5):338. doi:  
585 10.1037/h0025665.

- 586 16. Morgan D, Ozanne-Smith J. Surf bather drowning risk and exposure-related factors  
587 identified by an expert panel. *International Journal of Aquatic Research and*  
588 *Education*. 2012;6(4):336-49.
- 589 17. Edwards LK. Analysis of variance overview. In: Edwards LK, editor. *Applied*  
590 *Analysis of Variance in Behavioral Science*. New York: Marcel Dekker; 1993. p. 1-  
591 23.
- 592 18. Morgan D, Ozanne-Smith J, Triggs T. Self-reported water and drowning risk  
593 exposure at surf beaches. *Aust N Z J Public Health*. 2009;33(2):180-8. doi:  
594 10.1111/j.1753-6405.2009.00367.x.
- 595 19. Short AD. *Beaches of the Victorian Coast & Port Phillip Bay: A Guide to their*  
596 *Nature, Characteristics, Surf and Safety*. Sydney: Sydney University Press; 1996.
- 597 20. Bruine de Bruin W, Fischbeck PS, Stiber NA, Fischhoff B. What number is  
598 “fifty-fifty”? Redistributing excessive 50% responses in elicited probabilities. *Risk*  
599 *Anal*. 2002;22(4):713-23. doi: 10.1111/0272-4332.00063.
- 600 21. Tabachnick BG, Fidell LS. *Experimental Designs using ANOVA*. Belmont:  
601 Thomson; 2007.
- 602 22. Howell DC. *Statistical Methods for Psychology*. 5th ed. Belmont: Duxbury Press;  
603 2002.
- 604 23. Cotton JW. Latin Square Designs. In: Edwards LK, editor. *Applied Analysis of*  
605 *Variance in Behavioral Science*. New York: Marcel Dekker; 1993. p. 147-96.
- 606 24. Chi MTH. Laboratory methods for assessing experts’ and novices’ knowledge. In:  
607 Ericsson KA, Smith J, editors. *The Cambridge Handbook of Expertise and Expert*  
608 *Performance*. Cambridge: Cambridge University Press; 2006. p. 167-84.
- 609 25. Ericsson KA. An introduction to Cambridge handbook of expertise and expert  
610 performance: Its development, organization, and content. In: Ericsson KA, editor. *The*  
611 *Cambridge Handbook of Expertise and Expert Performance*. Cambridge: Cambridge  
612 University Press; 2006. p. 3-19.
- 613 26. Stevens JP. *Applied Multivariate Statistics for the Social Sciences*. Mahwah:  
614 Lawrence Erlbaum Associates; 2002.
- 615 27. Field AP. *Discovering Statistics using SPSS 2nd ed*. London: Sage; 2009.
- 616 28. Federer WT. *Experimental Design: Theory and Application*. New York: Macmillan;  
617 1955.
- 618 29. van der Kloot W. Univariate and multivariate analysis of variance of longitudinal  
619 data. In: Bijleveld CCJH, Kamp LJTvd, Mooijaart A, Kloot WAvd, Leeden Rvd,  
620 Burg Evd, editors. *Longitudinal Data Analysis: Designs, Models, & Methods*.  
621 London: Sage; 1998. p. 155-206.
- 622 30. SPSS. *Statistical software for Windows 16.0.0 ed*. Chicago: SPSS; 2007.
- 623 31. SISA-Bonferroni [Internet]. 2008 [cited 22 November 2008]. Available from:  
624 <http://www.quantitativeskills.com/sisa/>.
- 625 32. Microsoft Corporation. *Microsoft Excel. 15.0.4885.1000 ed* 2013
- 626 33. Wiggins N, Hoffman PJ, Taber T. Types of judges and cue utilization in judgments of  
627 intelligence. *J Pers Soc Psychol*. 1969;12(1):52-9. doi: 10.1037/h0027364.

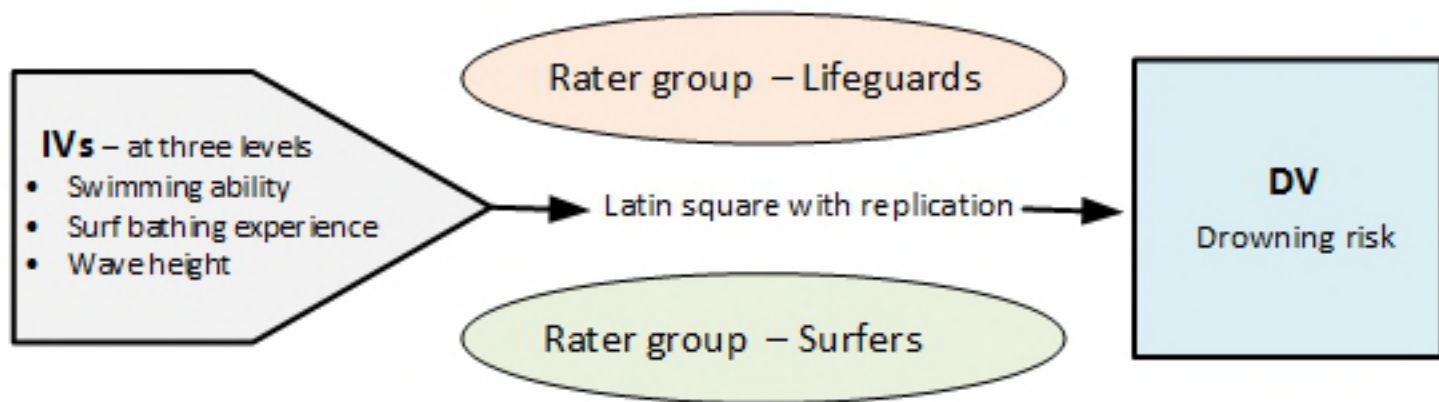
- 628 34. Green BF. Descriptions and explanations: A comment on papers by Hoffman and  
629 Edwards. In: Kleinmuntz B, editor. Formal Representation of Human Judgment. New  
630 York: Wiley; 1968. p. 91-8.
- 631 35. Doherty ME, Brehmer B. The paramorphic representation of clinical judgment: A  
632 thirty-year retrospective. In: Goldstein WM, Hogarth RM, editors. Research on  
633 Judgment and Decision Making: Currents, Connections, and Controversies New  
634 York: Cambridge University Press; 1997. p. 537-51.
- 635 36. Keselman HJ, Keselman JC. Analysis of repeated measurements. In: Edwards LK,  
636 editor. Applied Analysis of Variance in Behavioral Science. New York: Marcel  
637 Dekker; 1993. p. 105-45.
- 638 37. Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP. Common method biases in  
639 behavioral research: a critical review of the literature and recommended remedies. *J*  
640 *Appl Psychol.* 2003;88(5):879-903. doi: 10.1037/0021-9010.88.5.879.
- 641 38. Fisher RA. *The Design of Experiments.* 7th ed. Edinburgh: Oliver & Boyd; 1960.
- 642 39. McKay C, Brander RW, Goff J. Putting tourists in harms way—Coastal tourist parks  
643 and hazardous unpatrolled surf beaches in New South Wales, Australia. *Tourism*  
644 *Management.* 2014;45:71-84.
- 645 40. Queensland SLS. Gold Coast Beach Safe Program 2017. Available from:  
646 <http://lifesaving.com.au/beach-safety/goldcoastbeachsafe/>.
- 647 41. Western Australia SLS. Funding for Drone Surveillance Trial Announced [press  
648 release]. 10 October 2017 2016.
- 649 42. Life Saving Victoria. Nippers. Available from: [http://lsv.com.au/clubs-](http://lsv.com.au/clubs-members/nippers/)  
650 [members/nippers/](http://lsv.com.au/clubs-members/nippers/).
- 651 43. Surf Life Saving New South Wales. Surf Awareness & Safety Programs 2017.  
652 Available from: [http://www.surflifesaving.com.au/beach-safety/community-](http://www.surflifesaving.com.au/beach-safety/community-education/surf-awareness-safety-programs)  
653 [education/surf-awareness-safety-programs](http://www.surflifesaving.com.au/beach-safety/community-education/surf-awareness-safety-programs).
- 654 44. Cancer Council. SunSmart UV Alert 2009. Available from:  
655 <http://www.cancer.org.au/content/pdf/PreventingCancer/BeSunsmart/SunSmartUVale>  
656 [rtbrochure.pdf#\\_ga=1.85363865.1248960893.1485486072](http://www.cancer.org.au/content/pdf/PreventingCancer/BeSunsmart/SunSmartUVale).
- 657 45. Beachsafe. Welcome to Beachsafe 2017. Available from: <https://beachsafe.org.au/>.
- 658 46. Aschim H. Inflatable vests are no longer just for big wave pros. *Outside.* 2016 8  
659 February.
- 660 47. Maslakovic M. Wearable devices that keep your children safe. *Health & Fitness:*  
661 *Gadgets & Wearables.* 2016 22 November.

662

## 663 **Supporting information**

664 **S1 Table Scale Specification**

665 **S2 Table Questionnaire instructions**



*Note:* IV — independent variable; DV — dependent variable

**Fig. 1.** Study design overview

Fig 1

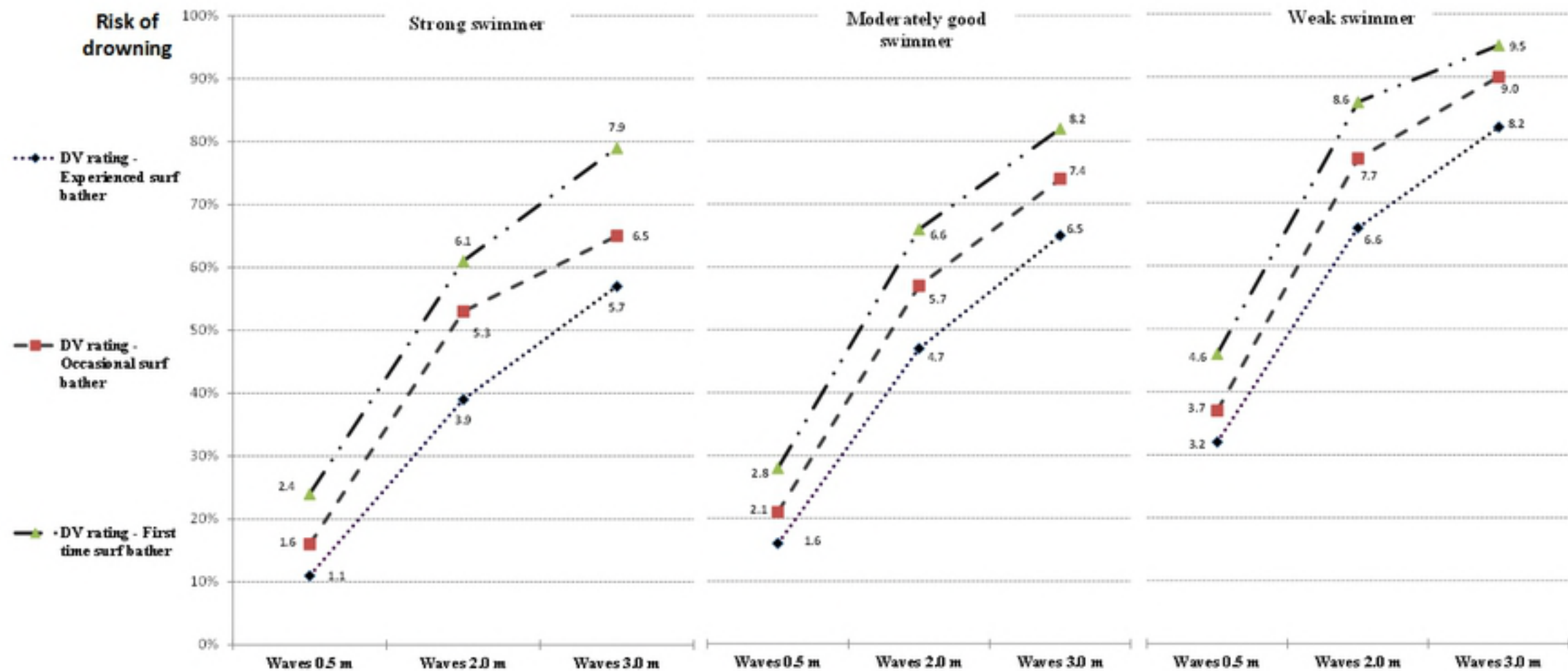
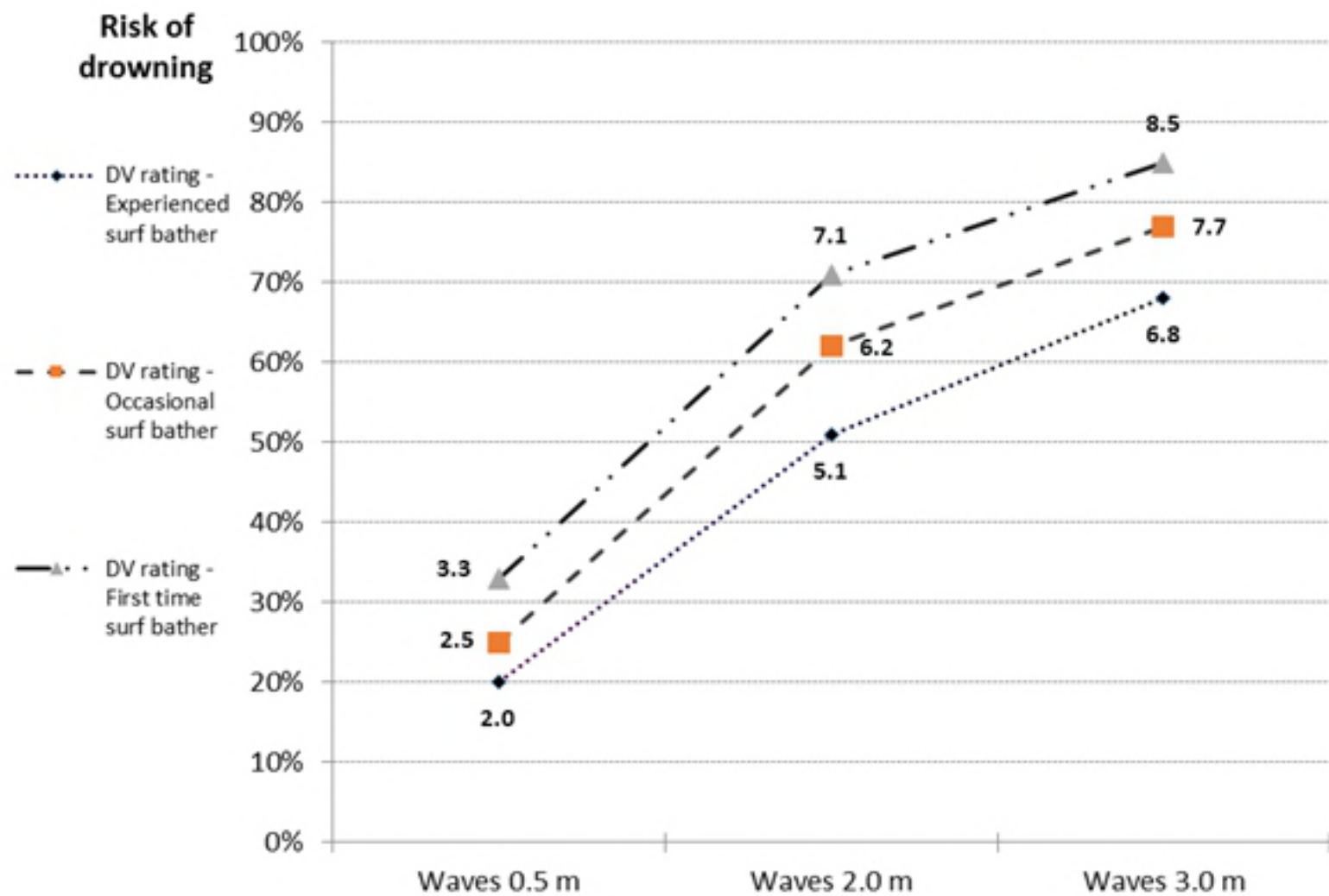


Fig. 2 Estimated marginal mean drowning risk scores on DV for swimming ability by surf bathing experience and wave height



**Fig. 3** Estimated marginal mean drowning risk scores for surf bathing experience by wave height

Fig 3