RESPONSE TO REVIEWERS

Please note that reviewers comments are in **bold**. These are followed by OUR REPLY and where our direct revision text is used, this text is in *italics*. All Line numbers in OUR REPLY refer to the revised version of the manuscript.

Reviewer to Reviewer #1 Comments:

This study provides an important evaluation of beach safety issues at unpatrolled beaches, which are commonly overlooked in Australia and other locations around the world. Because beach safety and the presence of lifeguards is not an important determinant on what beaches people visit, there is a need to develop beach safety knowledge and strategies at these sites. The paper is well-written and I only have a small number of suggested edits:

OUR REPLY:

We thank the reviewer for their positive comments about the study.

* It would be interesting to do a more thorough comparison of beach user responses from unpatrolled beaches (this study) to the responses from studies on patrolled beaches. Is there greater or lesser knowledge? The other studies are cited, but it would be interesting to compare the populations.

OUR REPLY:

This is a very good point and we have added comparisons in the Discussion in several places:

L362-363: 'Similar to the general Australian population (SLSA, 2021), 96% of respondents knew the meaning of the red and yellow flags on Australian beaches.'

L375-L386: 'These studies were all conducted on beaches with lifeguards, with the exception of Sherker et al. (2010), who did not discriminate results from respondents on patrolled and unpatrolled beaches. While direct comparisons of results of these previous studies are complicated due to survey bias, the different methods of lifeguarding and flag usage in other countries, varying beachgoer demographics and cultural factors, variable physical environmental conditions, as well as differences in the nature of the survey questions themselves, many similarities exist with the findings of this study. Sherker et al. (2010) found that 93% of Australian beachgoers, particularly males, felt that they could spot a rip current, but less than two-thirds actually could when shown photographs. Williamson et al. (2012) found that less than half of surveyed Australian beachgoers made the correct decision of where to swim when shown a photograph of a rip current. Of note, in a recent survey of beach safety knowledge of an Australian multi-cultural community, Woods et al. (2022) found that 83% of surveyed respondents had heard of the red and yellow flags, but only 29% said that they always swim between them, compared to 56% of the general Australian population (SLSA, 2021), and approximately half had either never heard of a rip current, or had, but did not know what they were.'

L388-398: 'In the United States, surveys of beachgoers at Pensacola Beach, Florida (Caldwell et al., 2013) found that 57% believed they could spot a rip current while only 15%

actually could and at Miami Beach, only 31% felt they could identify a rip current (Fallon et al.,2018). The images shown in Photo 2 and Photo 5 in Figure 6 were also used by Pitman et al. (2021) in a study of New Zealand beachgoers. In their study, 31% of beachgoers were able to identify a rip current in both photographs and 26% could identify a rip current in only one photograph. In comparison, 41% and 39.2% of beachgoers in this study correctly identified the rip current in Photo 2 and Photo 5 respectively (Table 3). The images shown in Photos 4 and 6 in Figure 6 have also been used to assess rip current identification of the Australian public in an ongoing national representative online survey by Surf Life Saving Australia (Lawes et al., 2021). According to unpublished SLSA data (J. Lawes, pers comm 27/1/2022) 56% and 35% of surveyed Australians between 2016 and 2020 correctly identified the rip current in Photos 4 and 6 respectively (Figure 6). These values are lower than those reported in this study with 71% and 40.3% of respondents correctly identifying the rip currents in those pictures (Table 3). However, not all of the surveyed Australians by SLSA were beachgoers.'

*How do the photographs used in Figure 5 compare to the beaches in which the study was completed, and how did the conditions in the photographs compare to the conditions on the days in which the surveys were conducted? Again how does this compare to previous studies of ability to spot rip currents?

OUR REPLY:

This is another important point. We have addressed this comment in several locations

L305: 'We have added text to the caption for Figure 5 (old Figure 4) noting that the image of Merry Beach was not taken during the survey data collection period.'

We have also created a new Supplement document (Supplement 1) that contains a Figure showing photographs taken at each beach during the survey period. These images capture the conditions at the time of the study. We refer to this Supplement in multiple locations in Section 2.1 and also address it with the following additional text:

L326-329 - It should be noted that none of the images in Figure 6 were of beaches involved in the survey. However, they represented rip currents present on NSW beaches that had similar morphological beach types and beach hazard ratings to those in the study and presented a range of surf conditions that were representative of conditions observed during the surveys (Supplement 1).

* A copy of the survey instrument is needed as an appendix or supplementary section.

OUR REPLY:

A copy of the survey was indeed submitted as a supplement file during the original submission process. With the addition of a new Supplement 1 in response to the reviewers comment above, the survey is now Supplement 2.

Response to Reviewer #2 (Sarah Trimble) Comments:

The introduction to this paper is thorough and clearly supports the two aims described. The paper is the first to address this subject, specifically, and provides needed insight within this body of literature. The methods are scientifically valid and thorough. All figures are necessary, with some minor edits or clarifications required.

OUR REPLY:

We thank the reviewer for these comments.

I recommend accepting the paper, after minor technical corrections and the improvement of one aspect: the only hiccup I found in the paper is in regards to the statistics used to show relationships (or lack thereof) between variables. First, the methods section 2.4 needs to be expanded to clarify (a) what is revealed by each statistical test and (b) therefore when each was used. At present, this section is merely a list. It needs to be expanded ever so slightly to that readers less familiar with these statistics than the authors can confidently understand the results.

OUR REPLY:

Thank you for pointing this out. We agree that it will be beneficial for the readers to have a brief description of each statistical test. We have therefore added text to the manuscript to describe when we have used each test depending on the variables of interest to this study and this research question.

L152-158: 'Kruskal-Wallis Test was used to determine if there were statistically significant differences between two or more ordinal variables (e.g. hazard perception vs. age group); Mann-Whitney U Test for differences between ordinal and binary variables (e.g. hazard perception vs. gender) and continuous and binary variables (e.g. rip score vs. gender); Chi-squared test for differences between categorical and binary variables (e.g. intention to go in water vs. gender); linear regression for relationships between continuous and ordinal variables (rip score vs. age group); binary logistic regression for relationships between binary and ordinal variables (e.g. intention go in water [yes/ no] vs. age group).'

Second, although the authors mention chi-square testing, they did not report it in several places where it is necessary. For example, within the paragraph beginning on line 301 is a description of the differences in each beach population's self-reported swimming ability. However, these values are within 10% of each other and when populations are this small (n= 104, 59, and 296) those differences may be explainable by random chance. A simple chi-square analysis between these answers can reveal whether the differences are significant, and either outcome would be worth reporting. It is important for the discussion to note whether the populations at each beach did have relatively the same swimming ability, or if more (or less) competent swimmers are more likely to visit one of the beaches. Similar edits are needed to lines 301-307, and 330-339.

OUR REPLY:

Thank you for your comment. Our response below also refers to similar comments received in the supplement pdf document review.

L303-305: These values are all quite close. Did you analyse with chi-square or a similar statistic? I suspect you would find that the range of these responses is within the realm of random chance, and not significant. If not significant, why report?

Thank you for raising this important point. We understand and agree that the previous wording may have led the wording to a different conclusion. We have originally included this sentence to describe the data set and never intended to statistically evaluate for differences between any variable and beach location.

Our pre-determined statistical aim was to understand how the variables age, gender, proximity to any coastal location, residency, swimming ability and swimming frequency influence someone's hazard perception, their evaluation of the importance of lifeguards, their intention to go in the water, their confidence in regard to rip spotting, their rip score and their ability to successfully provide a rip definition.

With the reviewer comments taken into consideration, we have reworded the sentence as follows:

L308-311: 'Respondents who indicated that they know or were unsure what a rip current was, were then asked to evaluate their own ability to visually spot a rip current with 44.7% being somewhat confident and 25.5% being very confident in their ability to do so. In terms of location, respondents were not at all confident or not very confident in similar proportions at the different beaches (Dreamtime Beach: 31.7%; Turimetta Beach: 25.4%, South Coast: 23.0%).'

L335-337: Again, percentages alone - especially those within 10% with fewer than 1000 subjects - do not indicate a trend or relationship. To clarify this finding, conduct a chi-square or similar test that will reveal whether there exists a difference in answers that is larger than possible with random chance.

We have conducted a Wilkoxon test that revealed that there is no statistically significant relationship between rip score and gender (p = .010). We understand that our wording could be misleading and thus have edited the sentence to:

L353-356: 'No difference existed between rip score and age (R^2 = .0251, F (4, 450) = 2.894, p = .022) or gender (W = 21784, p = .010, r = .108 – weak effect). Despite no statistically significant relationship between rip score and gender, males were slightly better at identifying rip currents with a mean score of 3.1 compared to females with 2.7.'

Specific comments are described in more detail below, and minor technical corrections to grammar and figures are included in the uploaded, commented PDF. Once addressed, I recommend this paper for publication in NHESS.

We thank the reviewer for their thoughtful and constructive comments and have addressed the General Comments above and now provide responses to the specific comments in italics directly below. Our responses to the commented supplement PDF begin on p. 8:

SPECIFIC COMMENTS

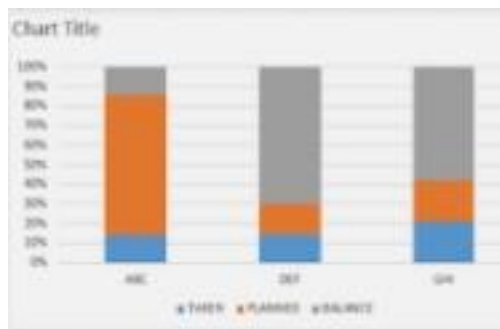
To my knowledge, all citations were appropriate. No major works were overlooked, and no inappropriate self-citation occurs. In addition to improvements required to the statistics (described above), here are a few more comments that need to be addressed, in the order in which they appear in the paper:

Based on responses to comments from Reviewer 1, we have added some additional new citations that we feel are appropriate to support our additions and were not published at time of our submission:

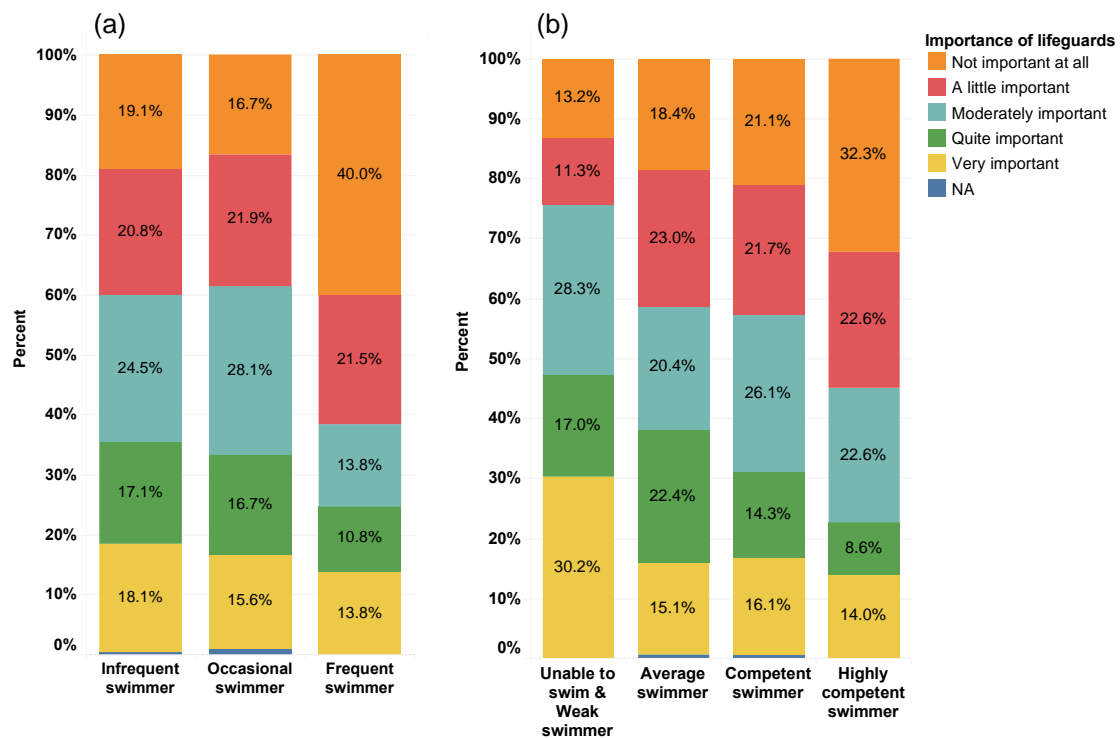
L395 – Lawes et al. 2021 (see L658 in References)

L384 – Woods et al., 2022 (see L743 in References)

We have also identified and corrected some mistakes in our citation format in the References

1) An easy-to-create bar chart would greatly aid in understanding the results reported in the paragraph beginning on line 228. At present, it is difficult to keep track of which subgroups are being described and compared, and a graphic would be a simpler way to differentiate these than adding excess explanatory text. Suggested addition: two "100% stacked" column charts showing the responses to "importance of lifeguards" (y-axis) where plot 1 shows distributions within swimming ability (x axis), and plot 2 is distribution within swimming frequency (x axis). Colors within each bar show popularity of each answer ('not important at all' through 'very important') as indicated by color within the bar. Example:

OUR REPLY:

Thank you for your suggestion and we agree that a bar chart visually helps to understand the results. We have created two bar charts based on your description and created a new Figure

(Figure 3) on L250

2) The paragraph beginning on line 253 needs correction. The third sentence states that "A significant difference" was found, but the statistics that follow appear to be similar, not different. All levels of self-reported swimming ability are described as having a similar proportion of respondents report that they do not perceive hazards. I do not know if it is a result of typos, or if increased explanation is needed. Either way, please address.

 OUR REPLY:

Thank you for your comment and we understand how this can be misleading. While a statistically significant relationship between hazard perception and swimming ability was found, the differences in proportions were quite small. We have edited the sentence as follows:

L255-260: While there was a statistically significant relationship between self-reported swimming ability and hazard perception, the effect size was weak and the differences in proportions small (H(4)=15.036, p = .005, f = .158 – weak effect).
All six people who self-reported to be unable to swim, rated the beach they were visiting as extremely hazardous (50%), very hazardous (16.7%) or somewhat hazardous (33.3%). In contrast, many (range: 48.7 – 57%) respondents from all other swimming ability categories rated the beach as not very or not at all hazardous.

3) The final column of Table 1 should refelct column "Total (n)" divided by 459, not 561. Because responses could be given more than one code, the total n=561 has no scientific meaning, but the proportion of respondents (459) who included a given response code is meaningful. Recalculate the final column, and leave empty the two rightmost cells in the bottom row.

We calculated the percentage based on the number of survey responses but understand that this is incorrect. We have divided the total n by 459 and edited the table to meet the formatting requirements. Furthermore, we made sure that the numbers in the manuscript are correct and have thus slightly changed the following paragraph.

L263-266: 'Responses were split into 18 separate codes with rip currents considered to be the primary hazard on all five beaches with 42.0% of all coded responses (Table 2). Respondents were also concerned about rough surf conditions (24.0%) at Dreamtime Beach, the rocks and cliffs at Turimetta Beach (28.8%) and the sun exposure at the South Coast (10.8%), although 13.2% of South Coast respondents had no concerns (Table 2).'

4) It is important to acknowledge the biases known to exist in this type of survey, such as conformity bias, agreement bias, and prestige bias. For example, with prestige bias, it is known that a respondent is more likely to give an answer that they believe the in-person surveyor will hold in higher regard. This can be accomplished with a few additional sentences somewhere within the paper. For example, in the paragraph beginning on line 311, where it is reported that people self-reported they were confident in their ability to swim were also confident in their ability to spot a rip. This result is currently presented as evidence that stronger swimmers are better at spotting rip, but without a swim test, it could just as easily be interpreted as "confident people are confident." If the authors feel it is more appropriate to address in the discussion than in results, a similar addition could be added around lines 345-350.

This is an important point and we have addressed this issue in the Discussion with the following text:

L371-374: 'It should be acknowledged that the results of this survey-based study are subject to a variety of well-known response biases that are inherent in many types of questionnaires, such as agreement bias, prestige bias and confirmation bias (Choi and Pak, 2005; Ménard et al., 2018; Davies, 2020). These types of bias also apply to previous studies involving surveys of beachgoer safety knowledge.'

This change involved adding two additional references:

L599 - Choi, B.C.K. and Pak, A.W.P.: A catalog of biases in questionnaires, Prev. Chronic Dis., 2(1), A13, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1323316/, 2005.

L615 - Davies, R. S.: Designing Surveys for Evaluations and Research. EdTech

Books. https://edtechbooks.org/designing_surveys, 2020.

5) In multiple places, it is necessary to cite known associations between gender and risk perception. The best location is lines 430-436. This will give the results in this paper important context. There is a large body of literature showing that females are more risk-averse than males in nearly all contexts and across all ages. Suggested citations: Gustafsod 1998, Savage 1993, DeJoy 1992.

OUR REPLY:

Thank you for your comment. We agree that it is important to point out the difference in risk perception between males and females and have therefore added the following sentence and citations:

L487-489: 'This difference in risk perception between gender is well documented in nearly all contexts and across all ages showing that males are more likely to undertake risky behaviour (DeJoy, 1992; Savage, 1993; Gustafson, 1998; Moran, 2011; Willcox-Pidgeon et al., 2018).'

The papers by Moran (2011) and Willcox-Pidgeon (2018) highlight this association in the context of drowning.

The Reference list has been updated with these new citations.

TECHNICAL CORRECTIONS

See commented PDF, attached.

Please note that the reviewers' comments are in **bold** and our replies follow below. We have indicated new text in *italics*. The line numbers in **bold** refer to the edited pdf version of the manuscript provided by the reviewer. The line numbers in italics refer to the revised manuscript.

L14: would prefer a quantitative value -- a percentage, or even just "majority" (if true)
Thank you for these comments and we have added percentages to make it clearer for the reader:

L14-16: 'Many unpatrolled beach users were infrequent beachgoers (64.9%) and weak swimmers (62.2%), with poor rip current hazard identification skills, who did not observe safety signage that was present, and yet intended to enter the water to swim (85.6%) despite being aware that no lifeguards were present.'

L19: Although I agree with this colloquial statement, scientifically we can't guarantee this with data. Instead, a scientific statement might be: "in recognition that many (the majority?) of Australian beaches will remain unpatrolled, yet still frequented, for the foreseeable future" ...or a less-wordy version of that.

Thank you for your comment and we agree that our wording was too colloquial. We have changed the sentence as follow:

L18-20: 'Future beach safety interventions in Australia need to extend beyond the standard 'swim between the flags' message in recognition that many of Australian beaches will remain unpatrolled, yet still frequented, for the foreseeable future.'

L20: This list is awkward and begs additional explanation. Suggest ending the sentence at "varied demographic groups of beach users."

Thanks for raising this point and we have added your suggestion into our manuscript.

L20-21: 'Future beach safety interventions for unpatrolled beaches should be tailored towards the varied demographic groups of beach users.'

L28: designates

We have fixed spelling.

L30-31: move period inside quotation mark and change these to double quotation marks

Thanks for pointing this out and we have changed it.

L30-31: 'The primary beach safety message promoted to beachgoers in Australia is to "always swim between the red and yellow flags."'

L37: are patrolled in Australia's most populous state, New South Wales (NSW)

We have added your suggestion into our manuscript.

L37: '... and only 20% of the 721 beaches are patrolled in Australia's most populous state, New South Wales (NSW) (SLSA, 2021).'

L39: ...that for more than 90% of coastal tourist parks in NSW, the closest beach is...

We agree with your suggestion.

L39-40: 'McKay et al. (2014) found that for more than 90% of coastal tourist parks in NSW, the closest beach is unpatrolled, or temporarily unpatrolled, and rated as 'hazardous' (Short and Hogan, 1994).'

L42: add citation for McKay et al?

We have added the citation for McKay et al. (2014) *(L42)*

L72-73: Was this known a priori? If so, cite sources -- or clarify by stating they have reputations as such. If these trends became apparent only after analyzing data collected for this study, then that should be stated and this sentence should be moved to results (not methods).

Thanks for raising this important point. We agree with the reviewer and feel that we have not emphasised this point enough in the manuscript. Yes, we have known the different reputations a priori, which was one of the reasons why we decided to conduct our research at those beaches. Dreamtime Beach has been ranked as one of the best beaches in Australia despite being a hazardous beach and many articles about it have appeared in Australian and

global mainstream and social media sources. We have attached two online links if you would like to read more about it. We have also used 'Beachsafe' to find out more about the beaches:

News.com.au (2018): https://www.news.com.au/travel/travel-updates/calls-to-stop-promoting-deadly-aussie-beach/news-story/5c0246a3ac9f2e502bbabf00667c28ce

McCutcheon (2018): https://www.abc.net.au/news/2018-02-22/concerns-after-dreamtime-beach-named-one-of-australias-best/9473244

SLS Beachsafe (n.d.) https://beachsafe.org.au/

We have therefore added text to the manuscript:

L73-74: 'The beaches have a reputation of showing different patterns of beach user demographics: ...'

L114-115: with whom they were visiting the beach; at which they were being surveyed - It is inappropriate to end a sentence with a preposition.

Thanks for pointing the mistakes out. We have changed the order of the preposition.

L119-122: 'Section 2 asked location specific questions, such as previous visitation to the beach at which they were being surveyed, reasons for visiting the beach, with whom they were visiting the beach, what activity they intended to participate in while at the beach, and if they would recommend the beach to others.'

Figure 1: preferably move to top of page 4 so that this Figure is located within 1 page of its first mention in line 72

We have moved Figure 1 to the end of page 4 after '2.1 Study site descriptions' and before '2.2 Survey design'.

L136-146: copy and paste error? remove repeated paragraph

Thanks for noticing such a big mistake. We have removed the repeated paragraph.

L162: Add: a brief description of when each statistic is the most appropriate test. Not all readers will be as familiar with these statistical tests as you are. Example: Chi-squared tests reveal when trends in categorical answers are beyond what may be expected by random chance.

Thanks for pointing this out. We agree that it will be beneficial for the readers to have a brief description of each statistical test. We have therefore added text to the manuscript to describe when we have used each test depending on the variables of interest to this study and this research question.

L152-158: 'Kruskal-Wallis Test was used to determine if there were statistically significant differences between two or more ordinal variables (e.g. hazard perception vs. age group); Mann-Whitney U Test for differences between ordinal and binary variables (e.g. hazard perception vs. gender) and continuous and binary variables (e.g. rip score vs. gender); Chi-

squared test for differences between categorical and binary variables (e.g. intention to go in water vs. gender); linear regression for relationships between continuous and ordinal variables (rip score vs. age group); binary logistic regression for relationships between binary and ordinal variables (e.g. intention go in water [yes/ no] vs. age group).'

L197-198: at which they were surveyed
We have changed the sentence.

L194-195: 'Respondents were asked to provide the main three reasons why they had chosen to visit the unpatrolled beach at which they were surveyed (Figure 2).'

L222-225: These two sentences appear to contradict. Is the first sentence only referring to males? Or perhaps despite the apparent differences described by the second sentence, the first is still true (no statistically significant relationship?)

We apologise that these sentences appear to be unclear and the latter is correct. Despite the apparent differences between age and the importance of lifeguards, we could not determine a statistically significant relationship between those two variables. Hence, we have adjusted the wording of these sentences.

L224-228: 'Older respondents valued the presence of lifeguards more with 41.4% of 45 – 59 year olds and 41.8% of those aged > 60 years rating it as quite important or very important, compared to only 24.3% of the 18-24 year age group. However, despite the apparent differences, no statistically significant relationship was found between age and the importance of lifeguards ($H^{(4)} = 11.244, p = .024, f = .128$ – weak effect).'

L228-233: Two simple bar charts would greatly aid in following these data.
y-axis: percent of group that gave a given answer
--- each bar goes to 100%
--- stacked colors within each bar show proportion of group that chose each option (i.e. not at all --> very important)
--- suggest using red & orange for "not important" and "a little"
x-axis: categories
--- chart 1: unable -->highly competent
--- chart 2: infrequent --> frequent
If my description is hard to follow, and you have access to Microsoft Excel, I am suggesting the option under Insert>Chart>2D Column>graphic where bars all go the maximum of y axis (100% stacked column)

Thanks for your suggestion. We have created the following bar chart and added it to the manuscript as a new Figure 3.

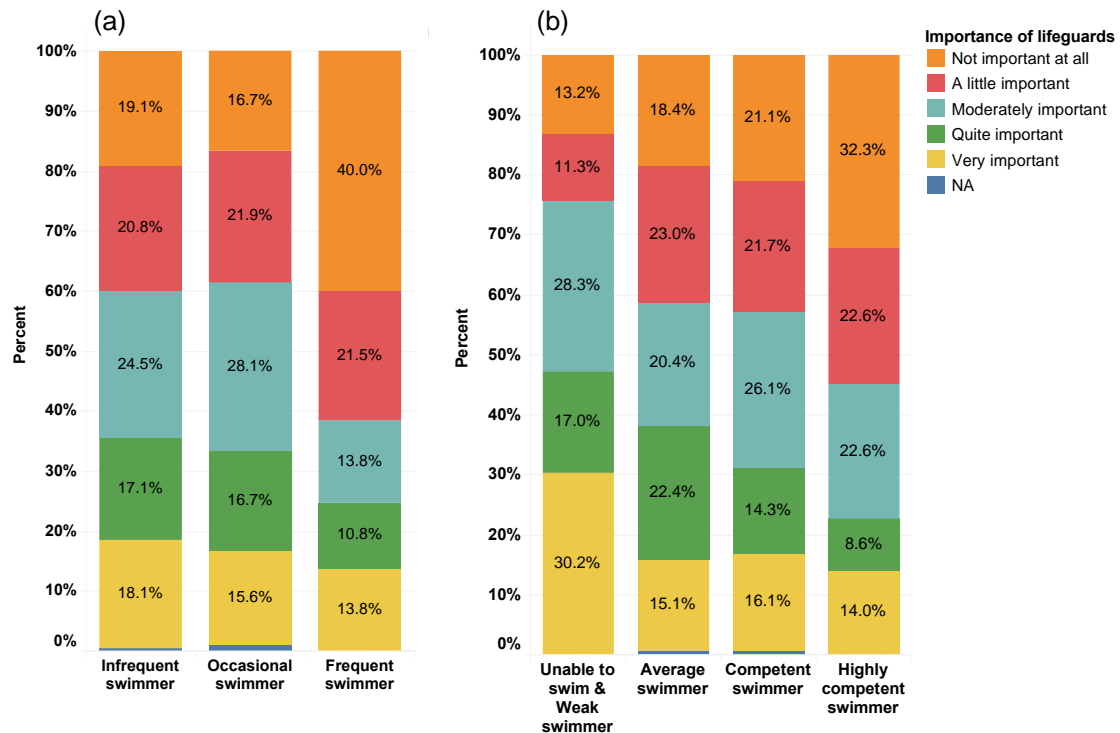



Table 1: add n=459; Please double check that this table meets the formatting requirements (i.e. border style) of NHESS

We have added the total number of responses and edited the table to meet the formatting requirements.

L243: suggest adding a brief sentence that repeats the Short 2007 hazard ratings from section 2.1, or incorporating them into each result description in this paragraph.

Thank you for your suggestion. We agree that it might be helpful to repeat the hazard ratings as all surveyed beaches have a high hazard rating. We have added the following sentence to the manuscript.

L238-239: 'Of note, according to Short (2007) the surveyed beaches are rated as moderately hazardous or as highly hazardous.'

L255: at that particular beach? or at any/all beaches?

The sentence starting in line 254-256 refers to all beaches. We have edited the sentence to make it clearer.

L253-255: 'Combining all beaches, almost one-third (29.4%) of respondents who lived greater than 51 km from the coast rated swimming and wading as not at all hazardous compared to only 9.4% of those who live between 1 – 9 km away from any coastal location.'

L256-259: This does not seem like a difference. All levels of self-reported competency had a low hazard perception? Please clarify.
Thank you for your comment and we understand how this can be misleading. While a statistically significant relationship between hazard perception and swimming ability was

found, the differences in proportions were quite small. We have edited the sentence as follows:

L255-260: While there was a statistically significant relationship between self-reported swimming ability and hazard perception, the effect size was weak and the differences in proportions small (H(4)=15.036, p = .005, f = .158 – weak effect).
All six people who self-reported to be unable to swim, rated the beach they were visiting as extremely hazardous (50.0%), very hazardous (16.7%) or somewhat hazardous (33.3%). In contrast, many (range: 48.7 – 57.0%) respondents from all other swimming ability categories rated the beach as not very or not at all hazardous.

Table 2: Determining the percentage with 561 is inappropriate because responses could receive more than one code. Instead, the total n at left should be divided by the total number of responses 459. You can then sort this column to show the true "popularity" of a given answer -- although the order will not change, the percentages will be an accurate reflection of how much of the surveyed population gave each answer.

Thank you for pointing out this mistake. We calculated the percentage based on the number of survey responses but understand that this is incorrect. We have divided the total n by 459 and edited the table to meet the formatting requirements. Furthermore, we made sure that the numbers in the manuscript are correct and have thus slightly changed the following paragraph.

L263-266: 'Responses were split into 18 separate codes with rip currents considered to be the primary hazard on all five beaches with 42.0% of all coded responses (Table 2). Respondents were also concerned about rough surf conditions (24.0%) at Dreamtime Beach, the rocks and cliffs at Turimetta Beach (28.8%) and the sun exposure at the South Coast (10.8%), although 13.2% of South Coast respondents had no concerns (Table 2).'

L280: Suggest inserting reminder: The south coast beaches were a, b, and c.

Thank you for your suggestion. We have inserted a reminder in the caption of Figure 4:
'The South Coast beaches were (a), (b) and (c).'

L298: no need to report n when all other statistics have only been reported as a %. These are two separate sentences. Also, "less" is used to refer to amounts that cannot be quantified (he was less hungry than before) and "fewer" is for quantifiable amounts. ...**into a rip current. Fewer people marked the rip...**

We have deleted the n value and adopted your changes.

L300-302: 'While most respondents (72.8%) chose a safe swimming spot away from the rip currents, almost a quarter (24.6%) would have entered the water directly into a rip current. Fewer people marked the rip current at Turimetta Beach (16.9%) compared to Dreamtime Beach (27.9%) and the South Coast (25.0%).'

L301: Add: how was this question posed to those who had said they did not know what a rip current was? Briefly explain.

We only asked respondents who answered that they knew or who were unsure what a rip current was to evaluate their own confidence. We made this decision to get a better understanding of their perception of their rip spotting ability.

L308-309: 'Respondents who indicated that they know or were unsure what a rip current was, were then asked to evaluate their own ability to visually spot a rip current ...'

L303-305: These values are all quite close. Did you analyse with chi-square or a similar statistic? I suspect you would find that the range of these responses is within the realm of random chance, and not significant. If not significant, why report?

Thank you for raising this important point. We understand and agree that the previous wording may have led the wording to a different conclusion. We have originally included this sentence to describe the data set and never intended to statistically evaluate for differences between any variable and beach location.

Our pre-determined statistical aim was to understand how the variables age, gender, proximity to any coastal location, residency, swimming ability and swimming frequency influence someone's hazard perception, their evaluation of the importance of lifeguards, their intention to go in the water, their confidence in regard to rip spotting, their rip score and their ability to successfully provide a rip definition.

With the reviewer comments, we have reworded the sentence as follows:

L308-311: Respondents who indicated that they know or were unsure what a rip current was, were then asked to evaluate their own ability to visually spot a rip current with 44.7% being somewhat confident and 25.5% being very confident in their ability to do so. In terms of location, respondents were not at all confident or not very confident in similar proportions at the different beaches (Dreamtime Beach: 31.7%; Turimetta Beach: 25.4%, South Coast: 23%).

L311: Self-reported ability, yes? There would likely be an association between people who are confident that they are good swimmers and confident that they can spot a rip, not because better swimmers are better at identifying rips but because some respondents want to exude confidence to the volunteer. Known biases: social desirability & conformity bias, prestige bias (likely here), and acquiescence of agreement bias.
"Prestige Bias is related to social desirability bias as it is based on an individual's personal desire to be seen in a positive light... For example, respondents may round up their income or report exaggerated amounts of time spent on worthy endeavors."
quick easy summary: https://edtechbooks.org/designing_surveys/validity_threats

We have addressed the issue of survey bias as describe on p. 7 of this document

L325: these are two separate statements

We have edited the sentences.
L332-334: 'The polygons were drawn based on the collective opinions of the authors, who have considerable experience in rip current identification. However, it is acknowledged that there is some degree of error involved in this process.'

L335-337: Again, percentages alone - especially those within 10% with fewer than 1000 subjects - do not indicate a trend or relationship. To clarify this finding, conduct a chi-square or similar test that will reveal whether there exists a difference in answers that is larger than possible with random chance.

We have conducted a Wilkoxon test that revealed that there is no statistically significant relationship between rip score and gender (p = .010). We understand that our wording could be misleading and thus have edited the sentence.

L353-356: 'No difference existed between rip score and age (R^2 = .0251, F (4, 450) = 2.894, p = .022) or gender (W = 21784, p = .010, r = .108 – weak effect). Despite no statistically significant relationship between rip score and gender, males were slightly better at identifying rip currents with a mean score of 3.1 compared to females with 2.7.'

L346-347: replace with or add the percentage; add clarifying & supportive statistic in parenthesis, such as: "XX% with a score of 4 or less & self-rated competent or highly competent swimmers"

We have added the percentages.

L367-369: 'While rip currents were the primary hazard that beachgoers were concerned about, many respondents were unable to correctly explain what a rip current was (10.2%), or identify them in photographs despite confidence in their ability to do so (38.5% of the somewhat confident and very confident respondents had a score of 2 or less).'

suggest switching the order of Figure 5 and Table 2, so that readers first see the questions (and answers) and see the results second. However, I do approve of the choice to show Figure 5 adjacent to table 1 in the results, rather than in the methods' description of the survey.

Thank you for your suggestion. These changes have been made. Please note, that Figure 5 (rip current images) is now labelled as Figure 6 in the manuscript.

Figure 6, Photo 2: couldn't this area be a channelised rip?

Figure 6, Photo 5: isn't this area farther down the beach also potentially a rip current?

Thank you for your comments and they highlight the difficulty of spotting rip currents in photos. Rip currents are often perceived as dark gaps and beachgoers are taught to look out for dark gaps. While this is accurate, not all dark gaps are rip currents. To address this issue in more detail and talk about the difficulties of spotting rip currents, we have added the following paragraph in the Discussion:

L400-408: 'It should also be noted that there are inherent limitations involved in assessing rip current identification skills using still images. There are issues with the angle and perspective from which the photographs were taken in relation to real-world perspectives of beachgoers on the beach as well as ambiguity of the perceived appearance of rip currents and non-rip current areas. For example, while the polygons denoting the rip current locations in Figure 6 are correct, it is possible that the gap in breaking waves below the polygon shown in Photo 2 (Figure 6b) could be construed as a channel rip current by a

respondent, even though it is a sandbar. Similarly, the region of dark water above the polygon in Photo 6 (Figure 6e) could be interpreted as a rip current, when in fact it was not. These limitations have previously been identified by Hatfield et al. (2013), Ménard et al. (2018) and Pitman et al. (2021), who have all suggested that future rip identification skills should be assessed with the use of video.'

L394-397: findings also supported by Brannstrom et al. -- highly suggest adding citation

We agree and have added additional Brannstrom et al. (2015) citation. Brannstrom et al. (2015) found that almost 50% (47.8%) of respondents did not notice any warning signs.

L447-449: 'Similar results were reported by Brannstrom et al. (2015) with 47.8% of respondents not noticing any beach safety signs. While necessary, warning signs are not necessarily as effective as authorities may assume (Matthews et al., 2014; Brannstrom et al., 2015)…'

L399-402: Divide this run-on sentence. It should be at least 2, and could be 3 separate sentences.

We agree that this sentence is too long and have divided it into two.

L452-455: 'A recent beach safety technology conference identified new technology (Beach Safety Research Group, 2021), which may increase swimmer safety at unpatrolled beaches. The establishment of emergency beacon stations, whereby beachgoers can alert emergency services about an incident immediately and, in some cases, access a flotation device to assist in a bystander rescue, may save not only the life of the individual in distress, but also the bystander rescuer.'

L402-405: Both of these statements need additional support. How would posting a drone operator at a beach aid in safety beyond shark spotting? How would monitoring the number and location of beachgoers aid in safety? How would these suggestions work on an otherwise unpatrolled beach?

We have addressed this comment with some clarification and additional text:

L455-461:'Technology may also play a future role with the potential use of drones, which have primarily been used in beach safety in relation to shark surveillance (Claesson et al., 2017; Butcher et al., 2019; Ajgaonkar et al., 2020), but also have the potential to be used to identify rip current locations. However, in the case of unpatrolled beaches, there is a logistic constraint of requiring a trained and certified operator to be present. The installation of low-cost remote video cameras has been shown to have the potential to monitor the number and location of people on beaches in real-time (Guillén et al., 2008; Morgan, 2018; Provost et al., 2019) and could be used by authorities to monitor peak usage times on unpatrolled beaches, which would have implications for future lifeguard provision.'

L418: fix to one of two options:
... both swimming *and* identifying...

...*either* swimming or identifying...

L472-473: '...many beach users were regional day visitors who were largely not confident or competent at both swimming and identifying rip currents.'

L431: Needed here: a quick sentence reporting that a large body of literature shows females self-report a higher perception of risk than males in nearly all subjects

Thank you for your comment. We have added the following sentence to point out the difference in risk perception between males and females.

L487-489: 'This difference in risk perception between gender is well documented in nearly all contexts and across all ages showing that males are more likely to undertake risky behaviour (DeJoy, 1992; Savage, 1993; Gustafson, 1998; Moran, 2011; Willcox-Pidgeon et al., 2018).'

We have added the additional references to the Reference list

L437: this study does not include data from other countries and only includes 2 non-domestic beachgoers so this must be removed

We have removed 'and other countries' from the sentence.

L493-495: 'However, the findings of this study suggest that Australia must move beyond the 'swim between the flags' message and also directly target the safety of beachgoers who choose to swim on unpatrolled beaches as a high-risk demographic.'