PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Staying Ahead of the Curve: Navigating Changes and Maintaining Gains in Patient Safety Culture—A Mixed-Methods Study
AUTHORS	Titi, Maher; Baksh, Maram; Zubairi, Beena; Abdalla, Rawia; Alsaif, Faisal; Amer, Yasser; Jamal, Diana; El-Jardali, Fadi

VERSION 1 – REVIEW

REVIEWER	Mehmet Top Hacettepe University, Turkey
REVIEW RETURNED	12-Mar-2020

GENERAL COMMENTS	I have reviwed your manuscript and study. I have some suggestions and commnets for your stuyd. 1. What ere originalities of your study? Originalities of this study should be detailed in introduction section. 2. The gap in which the research fills in the literature is not discussed in detail. 3. Directions for future research? 4. Data analysis should be improved. Correalation and regression analysis? 5. Normal distrubtions tested?

REVIEWER	Gregory Stock University of Colorado Colorado Springs United States
REVIEW RETURNED	07-Apr-2020

GENERAL COMMENTS	General comments This paper examines changes in patient safety culture in a large hospital in Saudi Arabia using a mixed methods study employing survey and interview data. For the most part, the paper is well written and provides some insights into the culture in this hospital. I have some comments the authors should consider in revising the manuscript. Specific comments • The very different response rates for the first and second data collections are puzzling and somewhat troubling. Do you know the reason for this difference? • Is there a reason that the respondents were overwhelmingly female? It appears that nurses were disproportionately represented in the sample, or do nurses make up that approximate proportion of workers in the hospital? Is there some explanation or justification? • Several of the dimensions of the survey were statistically significantly different across different administrations of the survey. However, I do not think they are meaningfully different, at least not

	when they are expressed as mean values. Moreover, many of the dimensions show both increase and decrease from year to year, so it is difficult to discern a meaningful pattern of the culture over time. • The presentation of percent positive responses was much more insightful, and I think the authors should highlight the differences between dimensions as well as over time. There are some inconsistencies in these differences. For example, management support for patient safety had one of the highest proportions of positive responses, but non-punitive response to error had one of the lowest percentages of positive responses. I would expect these two dimensions to be similar. What would be an explanation for this difference? Do you think the low proportion of positive non-punitive response to error was related to the fact that almost half the responses indicated that no events had been reported? • Is there any evidence in the qualitative interview data that would explain the differences between dimensions? • How would you generalize the results of this survey to other countries or types of hospital? How do these results compare to results of the survey in other countries? I do not expect that you would collect data in other countries, but there are many other published studies that have administered the survey in other countries.

REVIEWER	M.D. Cooper BSMS Inc, USA
REVIEW RETURNED	28-May-2020

GENERAL COMMENTS	The authors have conducted multiple distributions of a safety climate measure (NOT CULTURE) over an 8 year period. The manuscript needs a lot of work to get it into the correct format for a scientific journal report. The work has two main problems. [1] the statistics in Table 3 do NOT correspond to the authors statements of statistical significance; and [2] The authors have not discussed any study limitations, or acknowledged and accounted for Common Method Variance in their survey instrument (All psychology and business journals are recommending rejection on that basis alone). With regards to statistics contained in Table 3, the authors state they have conducted statistical tests between the mean scores for each year of distribution by topic, and at the bottom of the table provide the years for which the scores by topic differ significantly. After a visual inspection of the means and SDs in Table 3, I am certain there are no significant differences. i.e. there are no major differences in mean scores across the four distributions, and where there is some difference the SD (spread of scores) takes in the other mean scores. Thus, there just cannot be any statistically significant different scores. I calculated some of these differences by hand to double check. If the Authors intend to re-submit this manuscript to any journal, I strongly urge them to re-run their statistics, perhaps using ANOVA with a post-hoc Scheffe test. They should report all means and SDs and F values, and actual significance. These doubts about the veracity of statistical significance meant I discontinued the review at this point. I have attached a copy of the manuscript pdf containing my

	comments and text highlights to help them – contact publisher should you wish to see it.

REVIEWER	Gheed AlSalem Quality, accreditation and patient safety directorate Ministry of Health Kuwait
REVIEW RETURNED	06-Jun-2020

GENERAL COMMENTS	For the whole manuscript: Please revise for grammars corrections Page 5, Line 17 “Limitations of study” One of the most striking limitation of this study is the low response rate in 2019. I recommend adding it to this section and explaining the probable reasons behind it in a later section. Another limitation worth mentioning is that the interviews might not represent the range of views of health care staff in that hospital because they are mainly directors. Those limitations may affect the generalizability of the findings. Introduction Page 6, Line 13-14 “ …pattern of preventable adverse events in a study of eight developing countries, showing…” Wilson and Michel’s results date back to 2012, which is 8 years ago. Is this still the case today? I recommend adding more recent literature confirming the above point. Page 6, Line 19-27 “… Recent evidence continues identifying adverse events as a serious global issue affecting patient safety. Therefore, healthcare organizations made patient safety a top priority by identifying and preventing potential harms and enhancing patient outcomes”. DiCuccio and Mardon reviews date back to 2015 and 2010. Can you identify empirical evidence to support this claim? Do you have any recent, empirical, evidence to support their position on this? Also, in healthcare, the effect of hospital safety climate/culture on patient and worker safety outcomes is not entirely clear yet. Further research is needed as there is a shortage in the literature understanding how safety culture or climate impacts outcomes. It would be very useful to introduce such a point as safety research has moved beyond measuring safety culture/climate towards linking the results with patient/workers/ systems outcomes. Page 6, Line 38 “…recognized that healthcare organizations around the world regardless of income status”. This statement is vague..I this what is meant here is economic status. Also, can you identify empirical evidence to support this claim?

	Page 6, Line 43-53 "..Assessment of patient safety culture is being used to provide information support to managers and health policymakers, it can also help organizations in understanding the present state of …or programs and track change over time". Again, Can you identify empirical evidence to support the above point?. There is a lot of national, regional and international evidence to support the above. Add references to the above points. Page 7, Line 41-42 ".. lack of "just culture"" This point needs to be defined and clarified for the reader with appropriate references Page 7, line 43 "..Influence patient safety culture …impact patient outcomes" This is a very strong claim/statement. What levels of evidence were they empirically reporting on to make such a statement? The mentioned studies (20 24 25) measured patient safety climate/culture perceptions but did not prove the impact on patient outcomes. HSOPSC outcome measures are self-reported measures but are not outcomes. Page 7, Line 52 "..feedback as predictors for improved survey outcomes." Please clarify which kind of feedback. is it feedback and communication to staff about errors and incidents that happened within the healthcare facility. Page 8, Line 5-6 "..Organization's culture still considered relatively difficult.." Again, another vague statement. Please clarify which culture you are referring to. Also It should be emphasized that Safety culture and safety climate are derivatives of organisational culture and climate and not the same concept. Page 8, Line 12-13 "..quantitative with a qualitative measure) will help in deepen knowledge To be precise, this is a multiple method approach as each of the methods used are complete on their own. Since it is common to find the terms "multiple method", "mixed methods", and "multimethod" designs used interchangeably in the literature, it is necessary to clarify what is meant by the term when used. Safety climate questionnaires assess surface level features of the underlying safety culture in terms of perceptions of the workforce of procedures and practices in their workplace that show the level of priority given to safety in relation to other goals of a certain organisation. To explore the complex concept of safety culture/climate in healthcare, a multimethod approach can be viewed as most appropriate. The use of multiple methods increases the robustness of results through the cross-validation or "triangulation" that is accomplished when various data types and sources join and are found to be congruent. In my opinion, a multi-method, triangulated approach including both

	quantitative and qualitative methods was adopted for the study to gain a deeper understanding of the underlying culture while offsetting the intrinsic weaknesses associated with each approach when used on its own. Page 8, Line 19 "..Given the paucity of available evidence." Another vague statement that needs to be clarified... paucity of evidence around what? Page 8, Line 44 "..we will cross check the results with a qualitative assessment of safety culture to identify areas for improvement. Using a combination of data sources and research methods" A vague statement, Also, the term cross check is not appropriate. what you are trying to achieve here is, I assume is, a cross-validation or triangulation of results as I mentioned above. Page 8, Line 47-52 "..we will identify organizational changes that may have contributed to changes in patient safety culture and provide recommendations on how to leverage on these opportunities to maintain the gains." Please clarify what you mean here? Methods Page 9, Line 39 "The study used two different tools; in phase I: the tool used was adapted from the.." Clarify what you mean by two tools and the phases as you stated phase I but no following phases. Page 9, Line 42-43 "The validity and reliability of the .." Cronbach's Alpha is a commonly used measure for assessing the internal consistency/reliability of a set of items and not validity. Validity uses completely different measures. Also, kindly add an explanation to what is meant by Cronbach's with the recommended reference range. Page 10, Line 5 "…outcome dimensions" Again, kindly note that those are two variables (single, self-reported) and not dimensions of safety. objective safety outcome data include (observational studies, hospital records, incident reporting systems). Self-reported measures can be employed to assess the relationships between safety constructs. Still, data on perceptions of patient safety climate and on patient safety outcomes were collected from one source. Consequently, common method bias, biased recall, and social desirability may potentially affect the results. Page 10, Line 37 "…outcome dimensions" Again, kindly note that those are two variables (single, self-reported) and not dimensions of safety. Page 11, Line 10 "…[(number of positive response/total number of respondents on the item) × 100%]."

	Kindly correct the equation as (number of positive response/total number of all responses on the item) × 100% Page 11, Line 14-15. "…whereas those scoring less were considered areas for improvement." According to the original developers of the HSOPSC tool, the AHRQ defined patient safety areas in need of improvement as those survey dimensions where about 50% or more respondents answered negatively (strongly disagree/disagree) or "neither" to positively worded items, or 50% or more agreed (strongly agree/agree) with negatively worded items (Sorra and Nieva, 2004). Page 11, Line 33 "…interview participants were asked the same questions." What do you mean by this point. Also, did you use interviews and focus-groups. This point needs to be clearly stated. Kindly add the interview schedule as an appendix to provide clarity around the questions and the points of discussion during this phase of research Page 12, Line 17 "…the questionnaire did not collect any information that.." Which questionnaire? Page 12, Line 20-26 "…An email was sent for interviewees, including all the information about the voluntary nature of participation and explaining the entire study phase as well as a request to record the interviews. All interviews were carried out in a private area. Participant confidentiality was maintained at every stage of the study. " Please remove the repetition as it was mentioned before this point. Page 12, Line 30-31 "…The five stage 'framework approach' was employed for data analysis" It would be useful to add that the researchers followed a deductive approach for thematic data analysis. Page 12, Line 36 "…by the original research questions" There was no mention of any research questions so either add your questions before the study objectives or just delete the sentence. No mention of any ethical approval or ethics statement Results Page 13, Line 30 "…response rate of 59.8 % and 22.2 % respectively." Please add a reason to explain the big difference in the response rate between 2017 and 2019 Differences between respondents and non-respondents were not

	examined. Exploring that can add a deeper understanding. Also, there is no mention of the percentage of missing responses. The only mention was on Page 11, Line 10 "…Missing responses are excluded". I recommend adding it in Table 1 per section. it is common simply to drop these cases from analysis if the percentage is low. Page 14, Line 11 "…Frequency of events reported" and Table 1 Please correct this mistake. The measure here is called the NUMBER OF EVENTS REPORTED and not the frequency of events reported. Correct elsewhere. Page 17, Line 44 "which reflects that the interventions made at the hospital level have resulted in marked improvement in patient safety culture" Another statement that reads as unclear. I suggest adding a table or a list of the interventions that researchers are referring to with the years of implementation so that the reader can have a better understanding. Also, how can you state it as a marked improvement. Please add the level of significance in text and in table 3. Page 17, Line 37 "Comparing results from 2019-2012" It should be added that results must be interpreted with caution due to the low response rate. Page 17, Line 52 "…Teamwork within hospitals" Kindly correct the above as Teamwork within hospital units. Page 18, All Lines Explanation of results should be done in a chronological order and in a more organized way as you keep on going back and forth in time and it sounds confusing. Add subheadings if possible. Page 21, Line 43 "…and 13 departmental." Departmental what?? Directors/supervisors Please clarify Page 22, Line 7 "…documented observed changes" Clarify your point please. How were the changes documented. Please add a table that summarizes the main themes, sub-themes and main issues revealed at the sub-themes level during the interviews that were conducted. Page 23, Line 24 "Organizational changes since 2012" The slight dip in the composite scores in 2017 could be attributed to multiple organization-wide factors that occurred during this period. The shift of the human resources regulations and workflows from the

	traditional civil service to the self..” Please clarify this point further as I expect such change to lead to better findings. Page 23, Line 43 “paradigm shifts and escalations Please clarify what you mean here by paradigm shifts. Page 24, Line 43 “The first round of PSCS” Please put abbreviation in full. Page 25, Line 42 “documented suggestions” Again, were their interviews or a questionnaire with open comments that were analyzed. Please clarify as it is confusing. Page 26, Line 7 “Clearly, the changes in hospital structure led to..” Which changes in structure. Page 26, Line 27 “Sub-optimal staffing is potentially the most critical challenge...” This is a very authoritative statement. What levels of evidence were (references 40-41) citing / empirically reporting on to make such a statement? Would observational studies, investigating rates of error across a variety of countries (developed and developing) be more appropriate? Page 26, Line 32 “as they have an undeniable impact on patient care outcomes...” Can you identify empirical evidence to support this claim? Page 26, Line 42 “learning opportunities and system improvement” Please correct: Opportunities for learning Discussion and conclusion What are you contributing that is different to what has been done in many studies conducted in Saudi Arabia, other than updating prior results? If issues still exist regarding safety, as you point out, I would recommend doing something different in order to move past the limitations already mentioned. The current state of safety climate/culture, according to your results, offers valuable insights for developing and implementing effective strategies to promote improvements in patient safety areas including communication openness, error reporting and responses to errors. It would be helpful to list some of these strategies / interventions that you are aware of in a Table in the Discussion. This would support frontline managers operationalise how to ‘move to action’ for some of the insights surveys like this can provide. I also would expect a deeper analysis of results rather that stating the findings without critical analysis and understanding of what has been going on so far. Example: There was a discrepancy in the result of the outcome measure “Frequency of events reported” (70%) and the “Non-punitive response to errors” (39%). This may reflect the internal conflict that hospital staff have between their desire to report errors

	on one hand, and the predominant culture of blame and shame on the other hand. Also, it would be very useful to explore the dimensions based on nationalities with comparisons based on mean positive scores of the HSOPSC sub-scales.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1 (Mehmet Top)	
What are originalities of your study? Originalities of this study should be detailed in introduction section.	The originality of this study was highlighted in the end of the introduction as follows. We have edited to increase the clarity for the readership: Page 6: Given the paucity of evidence, the clear need for continuous monitoring of PSC and better understanding of how organizational culture changed after multiple interventions and milestones, we used a mixed method to assess PSC progressively over a period of eight years. To our knowledge this has not yet been conducted at this scale and will provide us with insight on key areas for improving PSC in healthcare organizations and persistent challenges in PSC that are difficult to change rapidly with simple interventions. The current study reports on organizational survey results for 2019 and also compares them to results of three previous assessments within the same multi-site facility (2012,(El-Jardali et al., 2014) 2015 (Alswat et al., 2017) and 2017).
The gap in which the research fills in the literature is not discussed in detail.	The mentioned gap was already emphasized in the previous item. To our knowledge this has not yet been conducted at this scale and will provide us with insight on key areas for improving PSC in healthcare organizations and persistent challenges in PSC that are difficult to change rapidly with simple interventions.
Directions for future research?	Added as a separate section after discussion titled: Implications for patient safety research
Data analysis should be improved. Correlation and regression analysis?	We have revised the data analysis models covering main areas of concern regarding data analysis by the reviewers regarding the following issues: - We ran a non-parametric Kruskal–Wallis H test given that not all composite scores are normally distributed. We revised the methodology to include this. The results are not different than the initial model with Bonferroni correction but are more appropriate given that

	we can't ascertain normal distribution for all composites. As for the variables on number of events and patient safety grade, we used a Chi-Square test. Furthermore, we revised the methodology for qualitative interviews to thematic analysis as the number of responses was not sufficient to construct tables.
Normal distributions tested?	We ran a non-parametric Kruskal–Wallis H test given that not all composite scores are normally distributed. This was re-emphasized in the methods (page 9) under 'data analysis – quantitative' Descriptive statistics such as frequency and percentage were used for data summarization. Given that data is not normally distributed for all composites, an Independent Kruskal-Wallis H test was used to determine whether a significant difference exists between survey composites between the 2012, 2015, 2017 and 2019 datasets. Chi-square test was used to determine whether a significant difference exists between 2012 and 2019 for the questions on number of events and patient safety grade.
Reviewer: 2 (Gregory Stock)	
The very different response rates for the first and second data collections are puzzling and somewhat troubling. Do you know the reason for this difference?	Page 11 Yes. We further clarified it in the *Results* section: 'The discrepancy in the response rate was observed as we stopped the second data collection in 2019 as the organization was preparing for the final survey of dual accreditation by both national and international accreditation bodies. We were concerned about the potential biased or false positive responses that might arise during the deep engagement of all of the staff at all levels in these accreditation activities'. Page 4 This statement was added to the limitations section as well: 'The lower response rate of the 2019 assessment round was observed compared to previous years and as such results should be interpreted with caution..'
Is there a reason that the respondents were overwhelmingly female? It appears that nurses were disproportionately represented in the sample, or do nurses make up that approximate proportion of workers in the hospital? Is there some explanation or justification?	We added an explanation to the results: Most of respondents in all four assessment rounds were nurses (50.1%, 78.3%, 56%, and 61.7%). The majority of respondents were female in the four assessment rounds that could be explained as the approximate proportion of our nursing staff is about 27% of the total number of the employees including both clinical and non-clinical staff. Moreover, our nursing staff are predominantly females

	(87%).
Several of the dimensions of the survey were statistically significantly different across different administrations of the survey. However, I do not think they are meaningfully different, at least not when they are expressed as mean values. Moreover, many of the dimensions show both increase and decrease from year to year, so it is difficult to discern a meaningful pattern of the culture over time.	This comment was addressed while comparing 2012 to 2019 results (page 16): As denoted in Table 3, Kruskal-Wallis test was conducted to compare results from all four surveys showed. Findings showed that significant differences lie across the four surveys with the exception for the composites on Supervisor/manager expectations and actions promoting safety and Non punitive Response to Error. A visual comparison of the mean scores between 2012 and 2019 show a steady increase in composite scores. It is worth noting that the highest scores were observed for 2015 as compared to 2012. Non-punitive response to error and staffing remained the lowest scoring composites during the period from 2012 to 2019 but the difference is not statistically significant. The highest-ranking composite for all surveys were Organizational Learning-Continuous Improvement.
The presentation of percent positive responses was much more insightful, and I think the authors should highlight the differences between dimensions as well as over time. There are some inconsistencies in these differences. For example, management support for patient safety had one of the highest proportions of positive responses, but non-punitive response to error had one of the lowest percentages of positive responses. I would expect these two dimensions to be similar. What would be an explanation for this difference? Do you think the low proportion of positive non-punitive response to error was related to the fact that almost half the responses indicated that no events had been reported?	We were aware of these discrepancies and have addressed this in the discussion section (page 25): - We observed a persistent discrepancy between the result of the outcome measure “frequency of events reported” and the “non-punitive response to errors” throughout the four PSC surveys despite the high reported average percentage of “feedback and communication about errors”. This may reflect the residual internal conflict that hospital staff have between their desire to report errors on one hand, and the predominant culture of blame on the other hand. - A further gap was noticed as management support for patient safety had one of the highest proportions of positive responses, but non-punitive response to error had one of the lowest percentages of positive responses. A possible explanation would be that the low proportion of positive non-punitive response to error was related to the fact that almost half the responses indicated that no events had been reported as in Table 4. Moreover, communicating about and addressing safety issues between healthcare leaders and front liners and developing training programs to help them understand their role in the development of safety culture could fill this gap. (Quenon et al., 2020)
Is there any evidence in the qualitative interview data that would explain the differences between dimensions?	Yes, these differences between (areas for improvement) and (areas of strength) were identified from the during qualitative interviews and were highlighted and summarized in Table 5 under the sections of suggestions to improve patient safety) and (changes in PSC in the facility) respectively (page 23)
How would you generalize the results of this survey to other countries or types of hospital? How do these results compare	This was highlighted in the Discussion At the regional level, many scholars have evaluated PSC

to results of the survey in other countries? I do not expect that you would collect data in other countries, but there are many other published studies that have administered the survey in other countries.	in different settings. Similar to our results, their findings show that the areas of strength were 'teamwork within units' and 'organizational learning and continuous improvement'. On the other hand, the area for improvement were 'promoting non-punitive response to error', 'encouraging the openness of communication among healthcare professionals', and 'facilitating hospital handoffs and transitions process'. (Alahmadi, 2010; Alswat et al., 2017; Ammouri, Tailakh, Muliira, Geethakrishnan, & Al Kindi, 2015; Badr, 2017; El-Jardali et al., 2010; Elmontsri et al., 2017; Khater et al., 2015). At the international level, similar areas for improvement were identified in hospitals. (Famolaro et al., 2018; Nie et al., 2013; Okuyama, Galvao, & Silva, 2018)
Reviewer: 3 (M.D. Cooper)	
The authors have conducted multiple distributions of a safety climate measure (NOT CULTURE) over an 8-year period. The manuscript needs a lot of work to get it into the correct format for a scientific journal report.	I agree with the concept you are putting forward but there are two major debates here: (1) Despite the fact that '…Safety culture and safety climate are clearly derivatives of organisational culture and climate, safety climate is often used interchangeably with safety culture. Moreover, Ginsburg and Tregunno et al argue for the lack of clarity in defining the construct of safety culture and climate in addition to the construct of patient safety culture. Therefore, the creation of a universal model or definition of safety culture is not straightforward.' Furthermore, 'there is no clear consensus in the difference between the safety culture and safety climate in healthcare settings. (2) 'PSC and safety climate are derivatives of the organizational culture and climate. Furthermore, patient safety climate is 'based on perceptions shared by individual in the work context' while PSC 'refers to the underlying assumptions and values that guide behavior in organizations'. Despite these conceptual differences, there is no clear consensus in the difference between the safety culture and safety climate as the research and practice use these two constructs interchangeably. Therefore, our study focuses on PSC as defined above.' **References:** • Alsalem G, Bowie P, Morrison J. Assessing safety climate in acute hospital settings: a systematic review of the adequacy of the psychometric properties of survey measurement tools. *BMC Health Serv Res* 2018;18(1):353. doi: 10.1186/s12913-018-3167-x • Hogden A, Churruca K, Bierbaum M. Safety culture assessment in health care: a review of the literature on safety culture assessment modes. • Griffin, M. A. and M. Curcuruto (2016). "Safety Climate in Organizations." Annual Review of Organizational Psychology and Organizational Behavior 3(1): 191-212. (3) The AHRQ tool that we have utilized is a well-known

	validated tool that was designed to measure PS Culture and is widely used to date. **Reference:** • Sorra JS, Dyer N. Multilevel psychometric properties of the AHRQ hospital survey on patient safety culture. BMC health services research. 2010 Dec 1;10(1):199. [371 Citations in Google scholar]
The statistics in Table 3 do NOT correspond to the authors statements of statistical significance; [With regards to statistics contained in Table 3, the authors state they have conducted statistical tests between the mean scores for each year of distribution by topic, and at the bottom of the table provide the years for which the scores by topic differ significantly. After a visual inspection of the means and SDs in Table 3, I am certain there are no significant differences. i.e. there are no major differences in mean scores across the four distributions, and where there is some difference the SD (spread of scores) takes in the other mean scores. Thus, there just cannot be any statistically significant different scores. I calculated some of these differences by hand to double check.] If the Authors intend to re-submit this manuscript to any journal, I strongly urge them to re-run their statistics, perhaps using ANOVA with a post-hoc Scheffe test. They should report all means and SDs and F values, and actual significance.	We have revised the data analysis models covering main areas of concern regarding data analysis by the reviewers in regards to the following issues: - We ran a non-parametric Kruskal–Wallis H test given that not all composite scores are normally distributed. We revised the methodology to include this. The results are not different than the initial model with Bonferroni correction but are more appropriate given that we can't ascertain normal distribution for all composites. As for the variables on number of events and patient safety grade, we used a Chi-Square test.
The authors have not discussed any study limitations, or acknowledged and accounted for Common Method Variance in their survey instrument (All psychology and business journals are recommending rejection on that basis alone).	There is a clear section for 'STRENGTHS AND LIMITATIONS OF THIS STUDY' in page 4 just before the introduction. We have added the CMV within the aforementioned limitations section. Change et al reported 4 approaches have been recommended in the literature as methods that researchers should use to avoid or correct CMV. Our case actually correlated to the first proposed remedy viz. 'to avoid any potential CMV in the research design stage by using other sources of information for some of the key measures. In particular, if possible, the dependent variable should be constructed using information from different sources than the independent variables'. Since we have used Mixed Methods rather than simply applying the survey method alone that have contributed to minimize the

	occurrence of spurious correlations due to CMV. **Reference:** Chang, S., van Witteloostuijn, A. & Eden, L. From the Editors: Common method variance in international business research. J Int Bus Stud 41, 178–184 (2010). https://doi.org/10.1057/jibs.2009.88
Reviewer: 4 (Gheed AlSalem)	
Page 5, Line 17 “Limitations of study” One of the most striking limitation of this study is the low response rate in 2019. I recommend adding it to this section and explaining the probable reasons behind it in a later section.	Page 11 Yes. We further clarified it in *the Results* section: ‘The discrepancy in the response rate was observed as we stopped the second data collection in 2019 as the organization was preparing for the final survey of dual accreditation by both national and international accreditation bodies. We were concerned about the potential biased or false positive responses that might arise during the deep engagement of all of the staff at all levels in these accreditation activities’. Despite the low response rate in 2019, the sample size we have is more than the minimum sample size recommended by the AHRQ. Moreover, all of the three previous PSCSs 2012, 2015, and 2017 conducted at our institution showed a good response rate >50% (85.7%, 57.6%, and 59.8% respectively). In this study we are reflecting and interpreting the four PSCSs simultaneously. Therefore, we do not think that the generalizability of our results are at risk. Reference: Sorra J GL, Streagle S, et al. . AHRQ Hospital Survey on Patient Safety Culture: User’s Guide.: (Prepared by Westat, under Contract No. HHSA290201300003C). AHRQ Publication No. 15-0049-EF (Replaces 04-0041). Rockville, MD: Agency for Healthcare Research and Quality, 2016. (page 12) Page 4 This statement was added to the limitations section as well: ‘The lower response rate of the second data collection was observed during the preparation for the final survey of dual accreditation in the organization.’
Another limitation worth mentioning is that the interviews might not represent the range of views of health care staff in that hospital because they are mainly directors. Those limitations may affect the	Further details were added to the results section and we do not believe this will have a substantial impact on the generalizability. Results section: ‘A total of 31 health care professionals were interviewed. Among which 10 were executive leaders (organization-wide or medical city-wide), 8

generalizability of the findings.	hospital directors, and 13 departmental chairmen. The researcher conducted four focus groups with a total of 60 front liner staff. The majority of the focus groups comprised physicians, pharmacists, nurses, allied health technicians.'
Introduction Page 6, Line 13-14 " …pattern of preventable adverse events in a study of eight developing countries, showing…" Wilson and Michel's results date back to 2012, which is 8 years ago. Is this still the case today? I recommend adding more recent literature confirming the above point.	Although we retrieved an updated citation (Schwendimann, R., Blatter, C., Dhaini, S. et al. The occurrence, types, consequences and preventability of in-hospital adverse events – a scoping review. BMC Health Serv Res 18, 521 (2018). https://doi.org/10.1186/s12913-018-3335-z), we were obliged to edit and delete this paragraph as part of complying with the word count required by BMJ open in the authors' guidelines which in turn makes this comment not applicable anymore.
Introduction Page 6, Line 19-27 "… Recent evidence continues identifying adverse events as a serious global issue affecting patient safety. Therefore, healthcare organizations made patient safety a top priority by identifying and preventing potential harms and enhancing patient outcomes". DiCuccio and Mardon reviews date back to 2015 and 2010. Can you identify empirical evidence to support this claim? Do you have any recent, empirical, evidence to support their position on this? Also, in healthcare, the effect of hospital safety climate/culture on patient and worker safety outcomes is not entirely clear yet. Further research is needed as there is a shortage in the literature understanding how safety culture or climate impacts outcomes. It would be very useful to introduce such a point as safety research has moved beyond measuring safety culture/climate towards linking the results with patient/workers/ systems outcomes.	We have added two more recent citations to support our position Bates & Singh, 2018 and World Health Organization, 2017 (page 5) Thank you for suggestion this outstanding point regarding the gap in patient safety research. We have added it under an additional section for *implication for future research*. Further patient safety research is needed as there is a shortage in the literature understanding how hospital safety culture impacts patient and worker safety outcomes. Furthermore, the research is moving towards linking the results of measuring safety culture with patient, workers, and health systems outcomes (page 27).
Introduction Page 6, Line 38 "…recognized that healthcare organizations around the world	Yes, true. Editing done. References were already provided from low and high income countries: (Alswat et al., 2017; Elmontsri,

regardless of income status". This statement is vague..I this what is meant here is economic status. Also, can you identify empirical evidence to support this claim?	Almashrafi, Banarsee, & Majeed, 2017; Neto AV et al., 2017). They show that the studies were conducted in variable economic levels.
Introduction Page 6, Line 43-53 "..Assessment of patient safety culture is being used to provide information support to managers and health policymakers, it can also help organizations in understanding the present state of …or programs and track change over time". Again, Can you identify empirical evidence to support the above point?. There is a lot of national, regional and international evidence to support the above. Add references to the above points.	Indeed, we agree with your highlighted point. The included citations had already addressed this but we added two more citations to emphasize and confirm it. **References:** • Reis CT, Paiva SG, Sousa P. The patient safety culture: a systematic review by characteristics of Hospital Survey on Patient Safety Culture dimensions. *Int J Qual Health Care* 2018 doi: 10.1093/intqhc/mzy171 • Bates DW, Singh H. Two decades since to err is human: an assessment of progress and emerging priorities in patient safety. Health Affairs 2018;37(11):1736-43.
Introduction Page 7, Line 41-42 ".. lack of "just culture"" This point needs to be defined and clarified for the reader with appropriate references	Rephrased to Moreover, lack of "just culture", that emphasizes shared accountability between leaders and staff to support error disclosure and organizational learning from mistakes,(Vogelsmeier, Scott-Cawiezell, Miller, & Griffith, 2010) poor communication, and shortage of staff negatively influence PSC . References: • Vogelsmeier, A., et al. (2010). "Influencing leadership perceptions of patient safety through just culture training." J Nurs Care Qual 25(4): 288-294.
Introduction Page 7, line 43 "..Influence patient safety culture …impact patient outcomes" This is a very strong claim/statement. What levels of evidence were they empirically reporting on to make such a statement? The mentioned studies (20 24 25) measured patient safety climate/culture perceptions but did not prove the impact on patient outcomes. HSOPSC outcome measures are self-reported measures but are not outcomes.	Yes, true. Editing done. Moreover, lack of "just culture", that emphasizes shared accountability between leaders and staff to support error disclosure and organizational learning from mistakes,(Vogelsmeier, Scott-Cawiezell, Miller, & Griffith, 2010) poor communication, and shortage of staff negatively influence PSC .
Introduction Page 7, Line 52 "..feedback as predictors for improved survey outcomes." Please clarify which kind of feedback. is it feedback and communication to staff	Editing done. …… organizational learning and feedback and communication about error as predictors for improved survey outcomes.

about errors and incidents that happened within the healthcare facility.	
Introduction Page 8, Line 5-6 "..Organization's culture still considered relatively difficult.." Again, another vague statement. Please clarify which culture you are referring to. Also It should be emphasized that Safety culture and safety climate are derivatives of organisational culture and climate and not the same concept.	Editing done. Measuring the PSC, is still considered relatively difficult and is well documented in the literature. (Gershon et al., 2004; Gutberg & Berta, 2017). And further explanation was added to the introduction
Introduction Page 8, Line 12-13 "..quantitative with a qualitative measure) will help in deepen knowledge To be precise, this is a multiple method approach as each of the methods used are complete on their own. Since it is common to find the terms "multiple method", "mixed methods", and "multimethod" designs used interchangeably in the literature, it is necessary to clarify what is meant by the term when used. Safety climate questionnaires assess surface level features of the underlying safety culture in terms of perceptions of the workforce of procedures and practices in their workplace that show the level of priority given to safety in relation to other goals of a certain organisation. To explore the complex concept of safety culture/climate in healthcare, a multimethod approach can be viewed as most appropriate. The use of multiple methods increases the robustness of results through the cross-validation or "triangulation" that is accomplished when various data types and sources join and are found to be congruent. In my opinion, a multi-method, triangulated approach including both quantitative and qualitative methods was adopted for the study to gain a deeper understanding of the underlying culture while offsetting the intrinsic weaknesses associated with each approach when used on its own.	We have elaborated on the mixed methods approach in the introduction: 'To explore the complex concept of PSC in healthcare, the mixed methods approach, by collecting and analyzing both quantitative and qualitative data, can be viewed as the most appropriate to increase the robustness of results and gain a deeper understanding of the underlying PSC while offsetting the intrinsic weaknesses associated with each approach when used on its own.' With an additional citation: **Reference:** • Shorten A, Smith J. Mixed methods research: expanding the evidence base. Evidence-Based Nursing 2017;20:74-75. Despite our agreement with the overall query and clarification of the review, we have concerns regarding using the 'triangulation' term that has been used in so many different ways that the editors of the Journal of Mixed Methods Research have recommended discontinuing the use of this term. **Reference:** • Fetters, Michael D., and José F. Molina-Azorin. "The journal of mixed methods research starts a new decade: principles for bringing in the new and divesting of the old language of the field." (2017): 3-10.

Introduction Page 8, Line 19 "..Given the paucity of available evidence." Another vague statement that needs to be clarified... paucity of evidence around what?	The required editing was applied to the introduction and methods. • 'To explore the complex concept of PSC in healthcare, the mixed methods approach, by collecting and analyzing both quantitative and qualitative data, can be viewed as the most appropriate to increase the robustness of results and gain a deeper'. understanding of the underlying PSC while offsetting the intrinsic weaknesses associated with each approach when used on its own • 'Given the clear need for continuous monitoring of PSC and better understanding of how organizational culture changed after multiple interventions and milestones, we used mixed method to assess PSC progressively over a period of eight years. To our knowledge this has not yet been conducted at this scale and will provide us with insight on key areas for improving PSC in healthcare organizations and persistent challenges in PSC that are difficult to change rapidly with simple interventions.' • 'In this study, a mixed methods research design was applied in order to achieve the most comprehensive understanding of the research problem possible. The study combined the use of quantitative surveys with qualitative interviews with hospital staff of different levels to achieve study objectives. Additionally, we conducted cross-validation with the results of previous surveys carried out in the same site in 2012, 2015 and 2017.'
Introduction Page 8, Line 44 "..we will cross check the results with a qualitative assessment of safety culture to identify areas for improvement. Using a combination of data sources and research methods" A vague statement, Also, the term cross check is not appropriate. what you are trying to achieve here is, I assume is, a cross-validation or triangulation of results as I mentioned above.	
Introduction Page 8, Line 47-52 "..we will identify organizational changes that may have contributed to changes in patient safety culture and provide recommendations on how to leverage on these opportunities to maintain the gains." Please clarify what you mean here?	
Methods Page 9, Line 39 "The study used two different tools; in phase I: the tool used was adapted from the.." Clarify what you mean by two tools and the phases as you stated phase I but no following phases.	Corrected
Methods Page 9, Line 42-43 "The validity and reliability of the .." Cronbach's Alpha is a commonly used measure for assessing the internal consistency/reliability of a set of items and not validity. Validity uses completely different measures. Also, kindly add an explanation to what is meant by Cronbach's with the recommended reference range.	Edited The internal consistency and reliability of the original English version had been tested with Cronbach alpha between 0.62 and 0.85.
Methods Page 10, Line 5 "…outcome dimensions" Again, kindly note that those are two variables (single, self-reported) and not	Edited. We agree with this improvement. The HSOPSC includes 42 items grouped into 12

dimensions of safety. objective safety outcome data include (observational studies, hospital records, incident reporting systems). Self-reported measures can be employed to assess the relationships between safety constructs. Still, data on perceptions of patient safety climate and on patient safety outcomes were collected from one source. Consequently, common method bias, biased recall, and social desirability may potentially affect the results.	composite measures or composites. In addition to the composites, the survey includes two questions that ask respondents to provide an overall grade on patient safety and to indicate the number of reported events over the past 12 months.
Methods Page 10, Line 37 "…outcome dimensions" Again, kindly note that those are two variables (single, self-reported) and not dimensions of safety.	Deleted
Methods Page 11, Line 10 "…[(number of positive response/total number of respondents on the item) × 100%]." Kindly correct the equation as (number of positive response/total number of all responses on the item) × 100%	Editing done
Methods Page 11, Line 14-15. "…whereas those scoring less were considered areas for improvement." According to the original developers of the HSOPSC tool, the AHRQ defined patient safety areas in need of improvement as those survey dimensions where about 50% or more respondents answered negatively (strongly disagree/disagree) or "neither" to positively worded items, or 50% or more agreed (strongly agree/agree) with negatively worded items (Sorra and Nieva, 2004).	Editing done with rationale. Our decision to consider our cutoff point to be 70% for patient safety areas in need for improvement was based on the results of our previous two PSCSs conducted in the same organization. (page 10). Even if we apply the cutoff of 50%, we will still have the same areas for improvement (viz. communication openness, staffing, non-punitive response to error) since all of them scored less than 50%.
Methods Page 11, Line 33 "…interview participants were asked the same questions." What do you mean by this point? Also, did you use interviews and focus-groups. This point needs to be clearly stated. Kindly add the interview schedule as an	Editing was done. The interview schedule was added as appendix

appendix to provide clarity around the questions and the points of discussion during this phase of research	
Methods Page 12, Line 17 "…the questionnaire did not collect any information that.." Which questionnaire?	Corrected to 'Interview Tool'
Methods Page 12, Line 20-26 "…An email was sent for interviewees, including all the information about the voluntary nature of participation and explaining the entire study phase as well as a request to record the interviews. All interviews were carried out in a private area. Participant confidentiality was maintained at every stage of the study. " Please remove the repetition as it was mentioned before this point.	Done
Methods Page 12, Line 30-31 "…The five stage 'framework approach' was employed for data analysis" It would be useful to add that the researchers followed a deductive approach for thematic data analysis.	Editing was done. Thematic analysis was conducted for data collected through interviews. Coding was initially conducted by breaking responses into similar concepts and ideas. This was done by two members of the research team. Minimal discrepancies were found after cross checking the work between the two team members. Disagreements were discussed until consensus was achieved. This was followed by axial coding, which comprised the organization of the emerging concepts into themes and categories.
Methods Page 12, Line 36 "…by the original research questions" There was no mention of any research questions so either add your questions before the study objectives or just delete the sentence.	Deleted
Methods No mention of any ethical approval or ethics statement	The ethical approval was written already at the end of the paper, according to the journal format, under Ethical approval, The study protocol and instruments were reviewed and approved (No. E19-4315) by the Institutional Review Board (IRB), College of Medicine, King Saud University.
Results	Explanation was added under the Results section

Page 13, Line 30 "…response rate of 59.8 % and 22.2 % respectively." Please add a reason to explain the big difference in the response rate between 2017 and 2019 Differences between respondents and non-respondents were not examined. Exploring that can add a deeper understanding. Also, there is no mention of the percentage of missing responses. The only mention was on Page 11, Line 10 "…Missing responses are excluded". I recommend adding it in Table 1 per section. it is common simply to drop these cases from analysis if the percentage is low.	(The discrepancy in the response rate was observed as we stopped the second data collection in 2019 as the organization was preparing for the final survey of dual accreditation by both national and international accreditation bodies. We were concerned about the potential biased or false positive responses that might arise during the deep engagement of all of the staff at all levels in these accreditation activities. Despite the low response rate in 2019, the sample size we have is more than the minimum sample size recommended by the AHRQ. Moreover, all of the three previous PSCSs 2012, 2015, and 2017 conducted at our institution showed a good response rate >50% (85.7%, 57.6%, and 59.8% respectively). In this study we are reflecting and interpreting the four PSCSs simultaneously.) Regarding the missing responses, the AHRQ recommends excluding displaying their percentages within the survey results as stated in the manuscript. But for transparency we have added their total numbers (Sorra 2016 Page 30). We have added all the suggested missing data to Table 1
Results Page 14, Line 11 "…Frequency of events reported" and Table 1 Please correct this mistake. The measure here is called the NUMBER OF EVENTS REPORTED and not the frequency of events reported. Correct elsewhere.	Editing done
Results Page 17, Line 44 "which reflects that the interventions made at the hospital level have resulted in marked improvement in patient safety culture" Another statement that reads as unclear. I suggest adding a table or a list of the interventions that researchers are referring to with the years of implementation so that the reader can have a better understanding. Also, how can you state it as a marked improvement. Please add the level of significance in text and in table 3.	This section was deleted as it is clearly addressed in the discussion section
Results Page 17, Line 37 "Comparing results from	This was already addressed in the discussion section in addition to the Strengths and Limitations of this Study

2019-2012” It should be added that results must be interpreted with caution due to the low response rate.	section.
Results Page 17, Line 52 “…Teamwork within hospitals” Kindly correct the above as Teamwork within hospital units.	This section was deleted as it is clearly addressed in the discussion section
Results Page 18, All Lines Explanation of results should be done in a chronological order and in a more organized way as you keep on going back and forth in time and it sounds confusing. Add subheadings if possible.	This section was deleted as it is clearly addressed in the discussion section
Results Page 21, Line 43 “…and 13 departmental.” Departmental what?? Directors/supervisors Please clarify	Edited A total of 91 health care professionals were interviewed. Among which 10 were executive leaders (corporate level), 8 hospital directors, and 13 departmental chairmen and directors. The researcher conducted four focus groups with a total of 60 front liner staff. The majority of the focus groups comprised physicians, pharmacists, nurses, allied health technicians.
Results Page 22, Line 7 “…documented observed changes” Clarify your point please. How were the changes documented? Please add a table that summarizes the main themes, sub-themes and main issues revealed at the sub-themes level during the interviews that were conducted.	Table 5 was added with the required
Results Page 23, Line 24 “Organizational changes since 2012” The slight dip in the composite scores in 2017 could be attributed to multiple organization-wide factors that occurred during this period. The shift of the human resources regulations and workflows from the traditional civil service to the self..” Please clarify this point further as I expect such	Editing was done for clarification. ‘Introducing the human resources regulations and workflows during the shift from the traditional civil service to the self-operation system, as a national trend throughout the Kingdom, had a significant impact on recruitment, re-contracting, and staff retention and turnover.’

change to lead to better findings.	
Results Page 23, Line 43 “paradigm shifts and escalations Please clarify what you mean here by paradigm shifts.	Editing was done for clarification. ‘Among the negative effects were overwhelming the staff with changes and escalations of improvement during a short period of time.’
Results Page 24, Line 43 “The first round of PSCS” Please put abbreviation in full.	PSC survey(s) was retained as the relevant abbreviation throughout the text
Results Page 25, Line 42 “documented suggestions” Again, were their interviews or a questionnaire with open comments that were analyzed. Please clarify as it is confusing.	Editing done ‘Some of the ‘interviewee’ suggestions for improving patient safety included…….’
Results Page 26, Line 7 “Clearly, the changes in hospital structure led to..” Which changes in structure.	Rephrased and clarified to be: ‘Clearly, the aforementioned organizational changes led to tangible….’ (discussion section)
Results Page 26, Line 27 “Sub-optimal staffing is potentially the most critical challenge...” This is a very authoritative statement. What levels of evidence were (references 40-41) citing / empirically reporting on to make such a statement? Would observational studies, investigating rates of error across a variety of countries (developed and developing) be more appropriate?	*Additional evidence to support our statement we added: -* • Aiken LH, Sloane DM, Bruyneel L, et al. Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. The lancet 2014;383(9931):1824-30. • Sturm H, Rieger MA, Martus P, et al. Do perceived working conditions and patient safety culture correlate with objective workload and patient outcomes: A cross-sectional explorative study from a German university hospital. PloS one 2019;14(1):e0209487.
Page 26, Line 32 “as they have an undeniable impact on patient care outcomes...” Can you identify empirical evidence to support this claim?	Yes, we added • Lee SE, Scott LD, Dahinten VS, et al. Safety culture, patient safety, and quality of care outcomes: A literature review. Western journal of nursing research 2019;41(2):279-304
Results Page 26, Line 42 “learning opportunities and system improvement” Please correct: Opportunities for learning	Corrected
Discussion and conclusion What are you contributing that is different to what has been done in many studies	Thank you for this interesting suggestion. We added Table (6) as a finale to the discussion titled : ‘Strategies or interventions to promote improvements in patient safety’.

conducted in Saudi Arabia, other than updating prior results? If issues still exist regarding safety, as you point out, I would recommend doing something different in order to move past the limitations already mentioned. The current state of safety climate/culture, according to your results, offers valuable insights for developing and implementing effective strategies to promote improvements in patient safety areas including communication openness, error reporting and responses to errors. It would be helpful to list some of these strategies / interventions that you are aware of in a Table in the Discussion. This would support frontline managers operationalise how to 'move to action' for some of the insights surveys like this can provide.	
I also would expect a deeper analysis of results rather that stating the findings without critical analysis and understanding of what has been going on so far. Example: There was a discrepancy in the result of the outcome measure "Frequency of events reported" (70%) and the "Non-punitive response to errors" (39%). This may reflect the internal conflict that hospital staff have between their desire to report errors on one hand, and the predominant culture of blame and shame on the other hand.	*We added this point to the discussion* 'We observed a persistent discrepancy between the result of the outcome measure "frequency of events reported" and the "non-punitive response to errors" throughout the four PSC surveys despite the high reported average percentage of "feedback and communication about errors". This may reflect the residual internal conflict that hospital staff have between their desire to report errors on one hand, and the predominant culture of blame on the other hand.'
Also, it would be very useful to explore the dimensions based on nationalities with comparisons based on mean positive scores of the HSOPSC sub-scales	This suggestion is out of the scope and objective of our current study.

VERSION 2 – REVIEW

REVIEWER	mehmet t turkey
REVIEW RETURNED	07-Sep-2020

GENERAL COMMENTS	This study is related to patient safety. I have some suggestions for your study. 1. Originalities of your study should be explianed in detail. 2. Limitations should be explainede detailed. 3. Praitical implications to nursing management should be detailed in conclusion section. 4. Research design should be detailed.

	5. Directions for future research should be detailed. 6. Sampling method should be explained in detail.

REVIEWER	Gheed Isalem Kuwait Ministry of Health Research and development
REVIEW RETURNED	24-Oct-2020

GENERAL COMMENTS	Kindly correct table 5 to table 2s as it is not in the main script and part of the supplementary data.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1 (Dr. Mehmet Top, Hacettepe University)	
1. Originalities of your study should be explained in detail.	**Already detailed in bullets 1 & 2 in strengths and limitations right after abstract (page 4)** • This is the first study to explore the results of a survey on patient safety culture in four consecutive rounds in Saudi Arabia. • This study covered almost all categories of healthcare workers, including managers and care providers.
2. Limitations should be explained detailed.	**Already detailed in bullets 4 & 5 in strengths and limitations right after abstract (page 4)** • A lower response rate was observed for the 2019 assessment and compared to previous years; therefore, the results of this study should be interpreted with caution. • To minimize the occurrence of spurious correlations due to common-method variance in the survey instrument, we have utilized a mixed-methods design rather than simply applying the survey method alone.
3. Practical implications to nursing management should be detailed in conclusion section.	**Why should the focus be on nurses, they do comprise half the sample, but the results were analyzed in aggregate with no specific analysis targeting nurses**
4. Research design should be detailed.	**It is very thorough**
5. Directions for future research should be detailed.	**Revisions made (page 27):** Repeated assessments of patient safety culture can provide unparalleled insight for hospital leaders into organizational changes resulting from quality improvement initiatives. Future research should link the results of patient safety culture assessments with patient, worker, and health system outcomes.
6. Sampling method should be explained in	We added a paragraph for clarity under

detail.	the Quantitative section of the methods: Participants: The survey randomly sampled staff targeting 50% (4500) of clinical and non-clinical employees similar to the previous two assessment rounds by El-Jardali et al and Alswat et al.
Reviewer: 2 (Dr. Gheed Alsalem, University of Glasgow, Ministry of health)	
Kindly correct table 5 to table 2s as it is not in the main script and part of the supplementary data.	It is already labelled as 2s table in the supplementary document. It was written ‘table 5’ in the previous response to the reviewer 2 (Dr. Gheed Alsalem) by mistake.