

## HOW TO CREATE A GOOD TABLE (AND AVOID COMMON MISTAKES)

Samir Kumar Praharaj

Associate Professor, Dept. of Psychiatry, Kasturba Medical College, Manipal University, Manipal.

Correspondence: Dept. of Psychiatry, Kasturba Medical College, Manipal, Karnataka. PIN – 576104 E-mail: samirpsyche@yahoo.co.in

Tables help coherent presentations of statistical inferences and allow easy comparisons of datasets, and could be invaluable aids when you intend to describe research findings in scientific papers. Done in a proper manner, they can summarize a large amount of data in an easy to grasp way, and help reduce your word count when you have more information to share than the journal allows.

### IMPORTANT BASIC RULES

Before making a table, decide whether tables are the best means to convey that particular data. For example, while tables are a good option for smaller data sets, especially if the precise data values need to be shown, when you have to show a huge amount of data, figures maybe a better option. Also, some information, like the one showed in **Table 1**, maybe better incorporated in the article text itself than in a table (To avoid confusion, references to the tables of this manuscript are printed in bold.).

**Table 1:** Anxiety symptoms in 50 depressed patients

Symptom	No. of patients
Palpitation	14
Dry mouth	12
Dyspnea	9
Tremor	7
Sweating	4
Paresthesia	1

Although an integral part of the manuscript, a table ought to be self-explanatory — i.e., the reader should be able to fully understand its contents without having to refer to the text. Also, once you have shared some information in a table, do not repeat that same information in the text too. However, some repetition is okay if it helps in highlighting a major finding or in providing additional explanation of any unexpected or interesting observations.

One should not submit output tables from statistical software like SPSS unmodified. Such programs produce multiple tables per statistical test, with unnecessary details that deserve deletion. Also, in the tables churned out by such software, the variable names for the X and Y axis will be pulled from the values entered by the researcher during data entry, and those terms, often some idiosyncratic abbreviations, may not make much sense for other readers. So such variable names too will have to be modified appropriately. Besides, while designing meaningful tables that highlights only one or two specific aspects of a study, one may have to use and compile information from many such “readymade” tables. Boldface, italics or different font types should be only sparingly used in tables, as they rarely improve readability. Remember: leaner tables are aesthetically better. For the same reason, your tables should not be very long, and should ideally fit within one printed page of the journal — It is not

Please cite this article as: Praharaj SK. How to create a good table (and avoid common mistakes). Kerala Journal of Psychiatry 2015; 28(2):219-24.

**Table 2:** Title of the table

	Heading 1			Heading 2 (Test Statistic)	Heading 3
	Subheading 1 <sup>†</sup>	Subheading 2 <sup>a</sup>	Subheading 3		
Row heading 1				X*	
Row heading 2					
Row heading 3 <sup>b</sup>					

*Note: General notes appear here, including definitions and abbreviations*

<sup>†</sup>Footnote 1, <sup>a</sup>Footnote 2, <sup>b</sup>Footnote 3

\*p-value

very comforting to read a table that extends to another page. At times, it may be possible to accommodate a large table in a single page if it is designed in landscape orientation. If needed, huge tables can also be split to two or more smaller ones.

Journals usually set a limit on the number of tables that can be allowed in a manuscript. Some journals, however, allow extra tables to be submitted as supplementary material, mostly as online content.

#### STRUCTURE OF A TABLE

Tables usually consist of:

- A title,
- A matrix of column headings (and their subheadings) and side (row) headings,
- The field (the rows and columns containing the data, i.e. the cells), and
- Explanatory notes (i.e. footnotes) (See **Table 2**).

The internal horizontal and vertical rules are usually not included, but the pattern varies across journals. Tables you initially submit can feature internal horizontal and vertical rules, and that may get modified during designing of the publication draft. Also, when a manuscript is submitted for publication, the tables are usually to be incorporated after the references, with each table placed on separate pages — the journal will align them with the text when the publication draft is designed.

#### TABLE NUMBER AND TITLE

All tables in a manuscript need to be numbered in chronological order and referred to within the text (Example: “as shown in Table 1,...” or “Table 2 shows...”). Most commonly in a clinical research paper, “Table 1” describes the sample characteristics. **Table 3** exemplifies how both continuous and categorical sociodemographic characteristics can be incorporated in a single table.

Title of a table is usually presented at the top. It should be concise and better not exceed two lines. At the same time, it should be complete too, and clearly describe the table’s contents. It should mention whether the table contains group comparisons, ‘correlation between variables’ or ‘sample description’, etc. The statistical test used too may be indicated (Example: “linear regression analysis showing...”). It is a good idea to include the sample size too, as in “sample characteristics (N=100)” or “group differences in intervention (n=40) and control (n=42) subjects” (Use uppercase ‘N’ to denote total sample, and lowercase ‘n’ for sample size of subgroups).

#### HEADINGS OF COLUMNS AND ROWS

Headings of columns should be short (two or three lines maximum). Number of subjects in each group should be clearly mentioned with

**Table 3:** Sample characteristics (N=60)

	Patients n=30	Controls n=30	t/ $\chi^2$	p
Age (in years), Mean (SD)	28.1 (8.4)	29.0 (7.9)	-0.41	.68
Education years, Mean (SD)	11.1 (4.3)	11.2 (3.9)	-0.17	.86
<i>Gender, n(%)</i>				
Male	22 (73.3)	25 (85.3)	0.88	.35
Female	18 (26.6)	5 (16.6)		
<i>Marital status, n(%)</i>				
Single	13 (43.3)	17 (56.6)	1.07	.30
Married	17 (56.6)	13 (43.3)		
<i>Socioeconomic status, n(%)</i>				
Low	xx (xx. x)	xx (xx. x)	x. xx	.xx
Middle	xx (xx. x)	xx (xx. x)		
High	xx (xx. x)	xx (xx. x)		
<i>Residence, n(%)</i>				
Rural	xx (xx. x)	xx (xx. x)	x. xx <sup>†</sup>	.03
Urban	xx (xx. x)	xx (xx. x)		
<i>Family type, n(%)</i>				
Nuclear	xx (xx. x)	xx (xx. x)	x. xx	.xx
Joint	xx (xx. x)	xx (xx. x)		

<sup>†</sup>Yates continuity corrected

**Table 4:** Group differences in clinical variables (N=133)

Variables	Group I n=45	Group II n=42	Group III n=46	F	p
BP (mmHg) <sup>†</sup>	43.2 (6.5)	xx. x (x. x)	xx. x (x. x)	5.4	<.001
BMI (Kg/m <sup>2</sup> )	20.1 (2.3)	xx. x (x. x)	xx. x (x. x)	xx. x	.354
C	xx. x (x. x)	xx. x (x. x)	xx. x (x. x)	xx. x	.221
D	xx. x (x. x)	xx. x (x. x)	xx. x (x. x)	xx. x	.945
E	xx. x (x. x)	xx. x (x. x)	xx. x (x. x)	xx. x	.061

Data are given as mean (SD) unless stated otherwise. BP: Blood pressure, BMI: Body-mass index. <sup>†</sup>Post hoc Tukey HSD showed significant difference between group I and II

**Table 5:** Pearson correlation (r) between WHO-QoL domain scores in bipolar patients (N=50)

	Domain 1	Domain 2	Domain 3	Domain 4	Domain 5
Domain 1	1				
Domain 2	0.46*	1			
Domain 3	0.23	0.69**	1		
Domain 4	0.76**	0.22	0.87**	1	
Domain 5	0.11	0.15	0.23	0.43*	1

\* $p < .05$ , \*\* $p < .01$  (two-tailed)

the group names. The test statistic and its value should be included, along with *exact significance values* if possible. (The format may vary across journals.) Sometimes, it is invaluable if separate columns are added for *effect size measures* or *post hoc tests* for multiple comparisons.

If the number of columns exceeds five or six, the table can be simplified by using symbols and shifting some information to footnotes. For example, in **Table 4** the post hoc test was relevant only for the first row, and was hence effectively shifted to the footnotes. Numbers that need to be compared with each other should be placed close together. If names of intervention and control groups are to be mentioned, they are placed from left to right.

In systematic reviews, depending on the subject of the review, the tables will have a different format. Commonly the headings include study (author name and year), design, participants (diagnosis and sample size), intervention, outcome, and remarks.

#### DATA FIELDS

The numbers are included in the cells, but their units of measurement (Example: ‘meq/L’ for lithium levels) are mentioned in the heading of the respective column or row. If it is necessary to add smaller or larger numbers, the unit of measurement can be altered and this too could be mentioned in the row headings (Example: For WBC counts, ‘ $\times 10^{12}/L$ ’ can be added).

For continuous variables, typically the mean and standard deviation (SD) are described. Sometimes however, especially if the data are skewed, median and interquartile range (IQR) are reported. It is

acceptable, and sometimes desirable, to report both. One common mistake is reporting standard error (SE), instead of SD, with the mean. As the SE values refer to sampling distribution, rather than the study sample, they would be smaller than SD and hence the readers could be misled.

If two values appear in a cell, they should be clearly described in the column or row heading [Example: “Mean (SD)” or “n (%)” or, in the footnote, “all values are expressed as mean (SD) unless stated otherwise”].

Usually the numbers in a cell are rounded off to two digits after decimal point. (Or even one digit is acceptable if precise values are not essential.) Whichever is used, there should be consistency in this regard throughout a manuscript.

For percentages, ensure that the numbers in a column add up to 100. Sometimes it may add up to 99.9 or 100.1 — this could be due to the rounding off, and do not require any correction, though if needed the discrepancy may be clarified in the footnote.

Some authors include an additional column or row just to mention the totals (row total or column total). This is usually not necessary and can clutter the table. Such values can be easily calculated by the readers, and are better avoided.

In some situations, there may not be any data in a cell (Example: Statistic column in Fisher’s exact test). Such *empty cells* may be left blank, or maybe filled with a horizontal dash. In tables containing correlation matrix, the cells containing repeated data may be left blank, as shown in **Table 5**.

## FOOTNOTES

All abbreviations (except standard ones) and symbols used in the table should be explained in the footnote. If a table is reproduced from another source, that too should be mentioned here.

Specific notes are conventionally indicated by superscript lowercase letters (e.g. a, b, c) and ordered from left to right and from top to bottom, starting with top left.

Footnote may also feature levels of statistical significance of the tests. (Common convention is to use asterisks: \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ). Reporting the *exact probabilities* ( $p = .02$ ) is always better than approximations ( $p < .05$ ). The way

significance is reported varies across journals (i.e. ‘P’ or ‘p’, ‘.05’ or ‘.05’, usual or bold), and details of a specific journal’s requirements can be gathered from its “instructions to authors” or from papers already published in it. Whatever be the case, one should use a consistent style throughout a manuscript.

## TABLE READY! BINGO

Once a table is ready, look if there are any surplus elements that can be removed (as shown in **Tables 6 and 7**). Double-check the data you entered, as mistakes are likely to have crept in — It is humiliating when a reader points out an obvious mistake later and you end up correcting it in a subsequent issue through an erratum.

**Table 6:** Comparison of schizophrenia patients (n=30), bipolar affective disorder patients (n=30) and normal control subjects (n=30) on expressed emotions (critical comments, hostility, dissatisfaction, emotional overinvolvement and warmth) using ANOVA statistic with post hoc Tukey HSD

Expressed emotion variable	Normal Control Subjects (A) n=30		Schizophrenia Patients (B) n=30		Bipolar Affective Disorder Patients (C) n=30		ANOVA F statistic	df	p	Effect size ( $\eta_p^2$ )	Post hoc Tukey HSD
	Mean	SD	Mean	SD	Mean	SD					
<i>Expressed emotion – Critical Comments</i>	6.321	1.526	8.416	1.121	7.935	1.354	4.512*	2,87	0.03	0.048	A<B A=C B=C
<i>Expressed emotion – Hostility</i>	10.012	2.143	11.431	1.346	11.124	1.852	0.413	2,87	0.420	0.002	A=B A=C B=C
<i>Expressed emotion – Dissatisfaction</i>	9.218	2.751	9.492	1.325	9.527	1.628	0.228	2,87	0.541	0.001	A=B A=C B=C
<i>Expressed emotion – Emotional over-involvement</i>	10.825	2.681	10.921	1.226	11.051	1.391	0.193	2,87	0.871	0.002	A=C A=B B=C
<i>Expressed emotion – Warmth</i>	7.881	1.891	8.182	1.291	7.914	1.641	0.381	2,87	0.513	0.003	A=B A=C B=C

\* $p < 0.05$  (two-tailed)

**Table 7:** Group differences in expressed emotions between patients with schizophrenia, bipolar disorder and normal controls (N=90)

Expressed emotion	Schizophrenia <sup>a</sup> n=30	Bipolar Disorder <sup>b</sup> n=30	Normal Controls <sup>c</sup> n=30	F (df=2,87)	p
Critical Comments <sup>†</sup>	8.4 (1.1)	7.9 (1.3)	6.3 (1.5)	4.5*	.03
Hostility	11.4 (1.3)	11.1 (1.8)	10.0 (2.1)	0.4	.42
Dissatisfaction	9.4 (1.3)	9.5 (1.6)	9.2 (2.7)	0.2	.54
Emotional over involvement	10.9 (1.2)	11.1 (1.4)	10.8 (2.7)	0.2	.87
Warmth	8.2 (1.3)	7.9 (1.6)	7.9 (1.9)	0.4	.51

Data are given as mean (SD) unless stated otherwise. \* $p < 0.05$  (two-tailed); <sup>†</sup> Effect size ( $\eta_p^2$ )=0.048, Post hoc Tukey HSD:  $a > c$

## TABLES IN ABSTRACTS AND POSTER PRESENTATIONS

In an abstract, tables are sometimes used (usually one) to highlight the most important findings. Specifically, when conference abstracts are published in journals, table can be helpful if they convey more information than what can be incorporated in the text. In such situations, it is recommended that the table should not exceed more than a third of the space of the abstract (some journals may allow a table separate from the text of abstract). The table must use the same font size as the abstract's text, or even smaller. Internal vertical and horizontal lines may be included if it improves readability of the table within the abstract.

Though posters commonly depend on illustrations to make their points, occasionally tables too are used. Large tables are better avoided, and font size must be large enough to be easily read from more than a four feet distance. Important findings can be highlighted or presented in boldface which enables drawing attention of readers. Color may be used to enhance the effect, but caution must be exercised to

not destroy the overall esthetics of the poster. During a power point presentation of study findings, tables are sparingly used. Larger tables published in journal papers are better split to smaller ones to maintain the font size and to highlight specific findings. Using boldface or different color may enable the important findings to stand out in the slide during oral presentation.

## FURTHER READING

1. American Psychological Association. Publication manual of the American Psychological Association. 6th ed. Washington (DC): American Psychological Association; 2010.
2. Durbin CG. Effective use of tables and figures in abstracts, presentations, and papers. *Respir Care* 2004; 49:1233–7.
3. Kotz D, Cals JW. Effective writing and publishing scientific papers, part VII: tables and figures. *J Clin Epidemiol* 2013; 66:1197.
4. Ng KH, Peh WC. Preparing effective tables. *Singapore Med J* 2009; 50:117-8.
5. Sternberg RJ. The psychologist's companion. A guide to scientific writing for students and researchers. 4th ed. Cambridge: Cambridge University Press; 2003.

Source of support: None Conflict of interest: None declared.