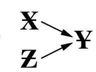
University of Waterloo W. H. Cherry

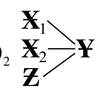
RELATIONSHIPS IN STATISTICS: Simpson's Paradox and Interaction

1. Background I – Illustrations of Simpson's Paradox [optional reading]

* **Lurking variate:** a non-focal explanatory variate (**Z**) whose differing distributions of values over groups of elements or units with different values of the focal variate, if taken into account, would meaningfully change an Answer about an **X**-**Y** relationship.
* **Confounding:** differing distributions of values of one or more *non*-focal explanatory variate(s) among two (or more) groups of elements or units [like (sub)populations or samples] with different values of the focal variate.
* **Comparison error:** for an Answer about an **X**-**Y** relationship that is based on comparing attributes of groups of elements or units with different values of the focal variate(s), comparison error is the difference from the *intended* (or *true*) state of affairs arising from:
  – differing distributions of lurking variate values between (or among) the groups of elements or units OR – confounding.

In other discussion (*e.g.*, in Statistical Highlights #57 and #63), the context for comparison error due to lurking variate(s)/confounding is comparative investigating of a *treatment* effect; the relevant *causal* structure (*e.g.*, from the upper half of page HL59.1 in Statistical Highlight #59), is case (8), shown at the upper right, with *focal* variate **X**, *respone* variate **Y** and lurking variate/confounder **Z**. In this Highlight #69, as summarized in the structure $(A)_2$ at the lower right, we broaden the discussion in two ways:

(8)


$(A)_2$


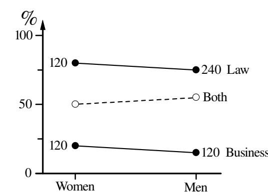
* we have two (or three) 'focal' variates [not necessarily all of equal interest in the Question context];
* we are *un*concerned with *causation* as the reason for the $\mathbf{X}_i$-**Y** and **Z**-**Y** associations, because the nature of the focal variates is such that we can*not* set their levels and this precludes using such focal variate(s) to manipulate the value of **Y**;
  – this is why the lower structure $(A)_2$ at the right above has *lines rather than arrows* between the variate symbols.

The data in Table HL69.1 below come from the discussion of Simpson's Paradox in Program 11 of *Against All Odds: Inside Statistics*; the context is possible sex discrimination in graduate admissions. Overall, the admission *rate* [or *proportion* (an *attribute*)] is *lower* for women (50% *vs.* 55% for men – see the bottom line of the Table) but, when the data are subdivided by school (Law and Business), the female admission rate is *higher* (by 5 percentage points) for *each* school. The (binary) response variate is school admission (Yes, No) and the lurking variate is women-to-men ratio among applicants; its effect is because:

* the two schools had appreciably *different* admission rates: 80 and 75% for Law, 20 and 15% for Business;
* *half* as many women as men (120 *vs.* 240) applied to Law but *equal* numbers of women and men (120) applied to Business.

Table HL69.1:

	WOMEN			MEN		
SCHOOL	Number of Applicants	ADMISSIONS Number	%	Number of Applicants	ADMISSIONS Number	%
Law	120	96	**80**	240	180	**75**
Business	120	24	**20**	120	18	**15**
Both	240	120	**50**	360	198	**55**



The diagram to the right of Table HL69.1 shows its data in graphical form; Simpson's Paradox is the *positive* slope of the middle dashed line for the data for *both* schools changing to a *negative* slope in the upper and lower lines for the schools *individually.*

In this illustration, the variates in the lower structure $(A)_2$ at the lower right above are:

$\mathbf{X}_1$ is an applicant's sex (female, male), $\mathbf{X}_2$ is the school applied to (Law, Business),

[In Tables HL69.5 and HL69.6 overleaf on page HL69.2, $\mathbf{X}_3$ is the level of study (Masters, Doctoral)],

Z is the (lurking variate) women-to-men ratio among applicants (discussed further in Section 2 overleaf on page HL69.2),

Y is the response to an applicant (admitted, not admitted). [On page HL69.2, **Y** is time for degree completion (minimum, longer).]

*Un*like investigating a treatment effect when there is more than one focal variate (*e.g.*, using a factorial treatment structure), the focal variate of primary interest in *this* Question context is $\mathbf{X}_1$, an applicant's sex.

The limitation imposed by lurking variates on an Answer to a Question about an **X**-**Y** relationship is illustrated further by the data in Tables HL69.2 to HL69.4; as the diagrams to the right of the tables emphasize, it is also possible to have:

Table HL69.2:

	WOMEN			MEN		
SCHOOL	Number of Applicants	ADMISSIONS Number	%	Number of Applicants	ADMISSIONS Number	%
Law	120	96	**80**	168	126	**75**
Business	120	24	**20**	120	18	**15**
Both	240	120	**50**	288	144	**50**

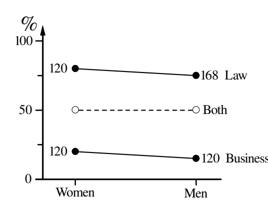
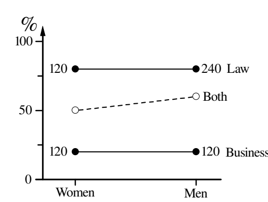


Table HL69.3:

	WOMEN			MEN		
SCHOOL	Number of Applicants	ADMISSIONS Number	%	Number of Applicants	ADMISSIONS Number	%
Law	120	96	**80**	240	192	**80**
Business	120	24	**20**	120	24	**20**
Both	240	120	**50**	360	216	**60**



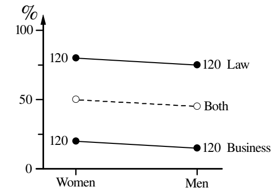
* the *same* overall admission rate for women and men but a *higher* rate for women in the two schools individually (Table HL69.2);
* a *lower* overall admission rate for women but the *same* rate for women and men in the two schools individually (Table HL69.3);
* a *higher* rate overall *and* in the two schools individually for women (Table HL69.4 overleaf on page HL69.2).

The effect of lurking variates on an **X**-**Y** relationship at a *second* level of subdivision is illustrated in Tables HL69.5 and

HL69.6 at the right below; a context for these data is the proportion of graduate students who complete their degree in the minimum time. In Table HL69.5, the proportion for women is *lower* overall, *higher* when subdivided by subject area (Law or Business) but again *lower* when subject area is subdivided by level (Masters or Doctoral). Similar effects are seen in Table HL69.6, except the proportions for women become *equal* when subdivided by subject area and *higher* when further subdivided by level.

Table HL69.4:

SCHOOL	WOMEN Number of Applicants	WOMEN ADMISSIONS Number	WOMEN ADMISSIONS %	MEN Number of Applicants	MEN ADMISSIONS Number	MEN ADMISSIONS %
Law	120	96	**80**	120	90	**75**
Business	120	24	**20**	120	18	**15**
Both	240	120	**50**	240	108	**45**



Probabilistically, subdividing is *conditioning* so that Tables HL69.1 to HL69.6, in illustrating Simpson's Paradox, show the limitation on an Answer which involves comparing *conditional* probabilities for a response variate with *different* conditionings; that is, comparing probabilities for **¥** given $\mathbf{X}_1$ and $\mathbf{X}_2$ with **¥** given only $\mathbf{X}_1$ (in Tables HL69.1 to HL69.4) or for **¥** given $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$ with **¥** given $\mathbf{X}_1$ and $\mathbf{X}_2$ or **¥** given only $\mathbf{X}_1$ (in Tables HL69.5 and HL69.6) – see the Section 3 discussion on the facing page HL69.3. Three further illustrative tables (like Table HL69.8 below) are discussed in Section 4 on pages HL69.3 and HL69.4.

Table HL69.5:

SCHOOL	WOMEN Number of Students	WOMEN COMPLETIONS Number	WOMEN COMPLETIONS %	MEN Number of Students	MEN COMPLETIONS Number	MEN COMPLETIONS %
Law: Masters	60	51	**85**	60	54	**90**
Doctoral	60	33	**55**	300	180	**60**
Bus.: Masters	60	27	**45**	20	10	**50**
Doctoral	60	9	**15**	100	20	**20**
Law	120	84	**70**	360	234	**65**
Business	120	36	**30**	120	30	**25**
Both	240	120	**50**	480	264	**55**

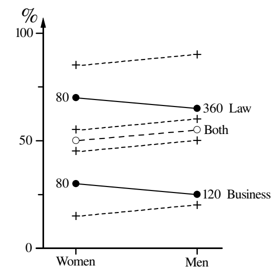
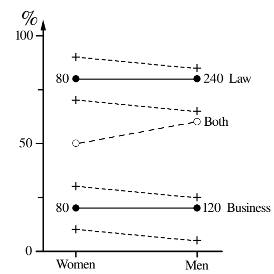


Table HL69.6:

SCHOOL	WOMEN Number of Students	WOMEN COMPLETIONS Number	WOMEN COMPLETIONS %	MEN Number of Students	MEN COMPLETIONS Number	MEN COMPLETIONS %
Law: Masters	60	54	**90**	240	204	**85**
Doctoral	60	42	**70**	80	52	**65**
Bus.: Masters	60	18	**30**	120	30	**25**
Doctoral	60	6	**10**	40	2	**5**
Law	120	96	**80**	320	256	**80**
Business	120	24	**20**	160	32	**20**
Both	240	120	**50**	480	288	**60**



The background in this Section 1, and in Sections 2 and 3 below and on the facing page HL69.3, are part of fuller discussion in Sections 3 to 7 on pages HL51.3 to HL51.6 in Statistical Highlight #51; readers unfamiliar with Highlight #51 may prefer its more detailed coverage of Simpson' Paradox and Interaction to that in this Highlight #69.

2. Background II – Reasons for Simpson's Paradox: population subgroups and weighted averages [optional reading]

The distorted calculation of the values of (population) attributes [like proportions (and averages)], which generates the 'paradox' illustrated in Section 1, is an instance of *weighted* combinations of the corresponding attributes of population *subgroups*. As shown in Table HL69.7 at the right, the attribute values in the last line of each of Tables HL69.1 to HL69.4 are weighted combinations of the attributes in the two table lines above them; what produces the changes in attribute values *relative* to each other is a change in *weights*. Each weight is determined by the (natural) *size* of a population subgroup; this size is the *lurking* variate whose change is responsible for the change in (the sign of) the **X**-**¥** relationship. The same idea applies to *each* of the *two* levels of subdivision in Tables HL69.5 and HL69.6. When the weights are *equal* (as in Table HL69.4), there is *no* 'paradox'.

Table HL69.7:

	Weighted percentage	Weights
Table HL69.1:	$\frac{120}{240}\times\mathbf{80} + \frac{120}{240}\times\mathbf{20} = \mathbf{50}$	$\frac{1}{2}$ $\frac{1}{2}$
	$\frac{240}{360}\times\mathbf{75} + \frac{120}{360}\times\mathbf{15} = \mathbf{55}$	$\frac{2}{3}$ $\frac{1}{3}$
Table HL69.2:	$\frac{120}{240}\times\mathbf{80} + \frac{120}{240}\times\mathbf{20} = \mathbf{50}$	$\frac{1}{2}$ $\frac{1}{2}$
	$\frac{168}{288}\times\mathbf{75} + \frac{120}{288}\times\mathbf{15} = \mathbf{50}$	$\frac{7}{12}$ $\frac{5}{12}$
Table HL69.3:	$\frac{120}{240}\times\mathbf{80} + \frac{120}{240}\times\mathbf{20} = \mathbf{50}$	$\frac{1}{2}$ $\frac{1}{2}$
	$\frac{240}{360}\times\mathbf{80} + \frac{120}{360}\times\mathbf{20} = \mathbf{60}$	$\frac{2}{3}$ $\frac{1}{3}$
Table HL69.4:	$\frac{120}{240}\times\mathbf{80} + \frac{120}{240}\times\mathbf{20} = \mathbf{50}$	$\frac{1}{2}$ $\frac{1}{2}$
	$\frac{120}{240}\times\mathbf{75} + \frac{120}{240}\times\mathbf{15} = \mathbf{45}$	$\frac{1}{2}$ $\frac{1}{2}$

3. Background III – Reasons for Simpson's Paradox: probability distributions [optional reading]

Table HL69.1 on page HL69.1 provides data which suggest an underlying probability function of a discrete trivariate distribution. To obtain this model, we first extend Table HL69.1 as in Table HL69.8 below to include three extra columns for 'Both sexes'. We then define five events and infer (approximate) values for ten probabilities – the vertical line means 'given that' in the eight *conditional* probabilities and $\cap$ denotes an *intersection* of events.

Table HL69.8:

SCHOOL	WOMEN Number of Applicants	WOMEN ADMISSIONS Number	WOMEN ADMISSIONS %	MEN Number of Applicants	MEN ADMISSIONS Number	MEN ADMISSIONS %	BOTH SEXES Number of Applicants	BOTH SEXES ADMISSIONS Number	BOTH SEXES ADMISSIONS %
Law	120	96	**80**	240	180	**75**	360	276	$\mathbf{76.\dot{6}}$
Business	120	24	**20**	120	18	**15**	240	42	**17.5**
Both schools	240	120	**50**	360	198	**55**	600	318	**53**

University of Waterloo W. H. Cherry

RELATIONSHIPS IN STATISTICS: Simpson's Paradox and Interaction (continued 1)

The (joint) trivariate model is shown in Table HL69.9 at the right below; summing its probabilities for one variate, we obtain the three (marginal) *bi*variate models in Tables HL69.10 to HL69.12. The smaller **bold** annotations in Tables HL69.9 to HL69.11 show how eight of the nine percentages in Table HL69.8 arise; for example, the 80% of women admitted to Law is $\frac{0.16}{0.2}$.

Event A:	Applicant is admitted ($\mathbf{Y}$ = yes; the complement $\bar{A}$ is $\mathbf{Y}$ = no)		
Event F:	Applicant is female ($\mathbf{X}_1$ = female)	Pr(F) = 0.4	Pr(A|F) = 0.5	Pr(A|F∩L) = 0.8
Event M:	Applicant is male ($\mathbf{X}_1$ = male)	Pr(M) = 0.6	Pr(A|M) = 0.55	Pr(A|F∩B) = 0.2
Event L:	Applicant applies to Law ($\mathbf{X}_2$ = Law)		Pr(A|L) = $0.7\dot{6}$	Pr(A|M∩L) = 0.75
Event B:	Applicant applies to Business ($\mathbf{X}_2$ = Business)		Pr(A|B) = 0.175	Pr(A|M∩B) = 0.15

Table HL69.9: Trivariate model for $\mathbf{Y}$, $\mathbf{X}_1$ and $\mathbf{X}_2$

	.....F.....		....M.....		
	L	B	L	B	
A	0.16	0.04	0.3	0.03	0.53
$\bar{A}$	**0.8** 0.04	**0.2** 0.16	**0.75** 0.1	**0.15** 0.17	0.47
	0.2	0.2	0.4	0.2	

Table HL69.10: Bivariate model for $\mathbf{Y}$ and $\mathbf{X}_1$

	F	M	
A	0.2	0.33	0.53
$\bar{A}$	**0.5** 0.2	**0.55** 0.27	0.47
	0.4	0.6	

We see that Table HL69.1 on page HL69.1 involves *parts* of the two multivariate distributions in Tables HL69.9 and HL69.10; it is therefore *un*surprising if comparisons among these parts, taken in isolation, yield seeming 'paradoxes'. It can be confusing that Table HL69.1 and those like it do not show *explicitly* percentages involving *complements* [like applicants '*not* admitted' (event $\bar{A}$)].

Table HL69.11: Bivariate model for $\mathbf{Y}$ and $\mathbf{X}_2$

	L	B	
A	0.46	0.07	0.53
$\bar{A}$	**$0.7\dot{6}$** 0.14	**0.175** 0.33	0.47
	0.6	0.4	

Table HL69.12: Bivariate model for $\mathbf{X}_1$ and $\mathbf{X}_2$

	L	B	
F	0.2	0.2	0.4
M	0.4	0.2	0.6
	0.6	0.4	

4. Simpson's Paradox and Interaction [The title matter of this Highlight #69.]

For convenience in this Section 4 (including labelling the three diagrams to the right of Tables HL69.13 to HL69.15 below and overleaf on page HL69.4), we retain the notation defined near the middle of page HL69.1:
$\mathbf{X}_1$ is an applicant's sex (female, male), $\mathbf{X}_2$ is the school applied to (Law, Business), $\mathbf{X}_3$ is the level of study [Masters, Doctoral],
$\mathbf{Y}$ is the response to an applicant (admitted, not admitted) or time for degree completion (minimum, longer),
$\bar{\mathbf{Y}}$ [the *average* of ($\mathbf{Y}$)] is the *percentage* of applicants admitted or who complete their degree in the minimum time.

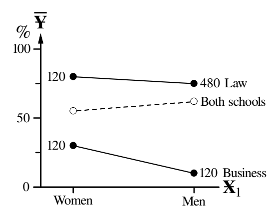
The diagrams illustrating Simpson's Paradox to the right of Tables HL69.1 to HL69.6 (on pages HL69.1 and HL69.2) are reminiscent of a diagram showing interaction (*e.g.*, in Note 5 overleaf at the bottom of page HL69.4); however, there are *differences*:

- ○ the Simpson's Paradox diagrams have an additional (dashed) line for the overall $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationship;
- ○ the instances of Simpson's Paradox in Tables HL69.1 to HL69.6 have only *parallel* (solid) lines for the $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationships for different values of $\mathbf{X}_2$ – that is, there is *no* interaction of $\mathbf{X}_1$ and $\mathbf{X}_2$ in their effects on $\mathbf{Y}$.

This restriction is *removed* in (another) reworking of Table HL69.1 and its diagram in Table HL69.13 below, where there *is* interaction of $\mathbf{X}_1$ and $\mathbf{X}_2$ in their effects on $\mathbf{Y}$ because the two solid lines in the diagram to the right of Table HL69.13 are *not* parallel.

Table HL69.13:

	..........WOMEN..........			..............MEN.............			.......BOTH SEXES......		
	Number of	ADMISSIONS		Number of	ADMISSIONS		Number of	ADMISSIONS	
SCHOOL	Applicants	Number	%	Applicants	Number	%	Applicants	Number	%
Law	120	96	**80**	480	360	**75**	600	456	**76**
Business	120	36	**30**	120	12	**10**	240	48	**20**
Both schools	240	132	**55**	600	372	**62**			



Thus, interaction *may* be involved in Simpson's Paradox but is not *required* for it to occur.

Earlier discussion at the upper left of page HL69.2, and preceding Table HL69.13, remind us that Simpson's Paradox and interaction *both* involve (approximate inferred) values of *conditional* probabilities for $\mathbf{Y}$, BUT:

- ○ Simpson's Paradox involves comparing these probabilities conditioned on two (or three) of the $\mathbf{X}$s with probabilities conditioned on one *fewer* (one or two) $\mathbf{X}$s; WHEREAS:
- ○ interaction is absent or present depending on the values of probabilities with the *same* conditioning on the $\mathbf{X}$s – these values determine whether the corresponding lines are or are not parallel.

NOTES: 1. Illustration of Simpson's Paradox from comparing *across* Tables HL69.1 to HL69.6 can overshadow comparisons *down* such tables. For example, in Table HL69.1 (reworked as Table HL69.8 on the facing page HL69.2), the six **bold** percentages for $\mathbf{X}_2$ (80 and 20, 75 and 15, $76.\dot{6}$ and 17.5) address a Question *different* from possible sex discrimination:

- ● *How do the admission standards of the Law and Business schools compare?*

The (hypothetical) data in Tables HL69.13 to HL69.15 above and overleaf on page HL69.4) indicate an appreciably *higher* admission standard for Business than for Law, unless the abilities in the two applicant pools are remarkably different.

2. A second reworking of Table HL69.1 and its diagram is given overleaf in Table HL69.14; as in Table HL69.13, there *is* interaction of $\mathbf{X}_1$ and $\mathbf{X}_2$ in their effects on $\mathbf{Y}$ but, for both schools combined, there is *no* sex difference in pro-

 (*continued overleaf*)

NOTES: (cont.) 2. portions, due to cancellation of effects in *opposite* directions for the schools individually.

Table HL69.14:

SCHOOL	WOMEN Number of Applicants	WOMEN ADMISSIONS Number	WOMEN ADMISSIONS %	MEN Number of Applicants	MEN ADMISSIONS Number	MEN ADMISSIONS %	BOTH SEXES Number of Applicants	BOTH SEXES ADMISSIONS Number	BOTH SEXES ADMISSIONS %
Law	120	96	**80**	180	153	**85**	300	249	**83**
Business	120	24	**20**	180	27	**15**	300	51	**17**
Both schools	240	120	**50**	360	180	**50**			

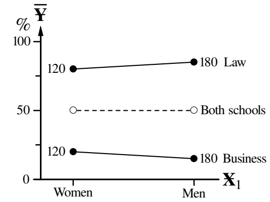
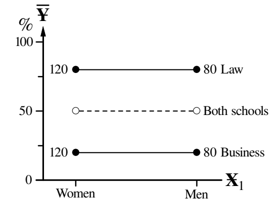


Table HL69.15 below and its diagram show, like Table HL69.14, *no* sex difference for both schools combined but this is now a consequence of the *individual schools* also showing this same behaviour – there is *no* interaction.

Table HL69.15:

SCHOOL	WOMEN Number of Applicants	WOMEN ADMISSIONS Number	WOMEN ADMISSIONS %	MEN Number of Applicants	MEN ADMISSIONS Number	MEN ADMISSIONS %	BOTH SEXES Number of Applicants	BOTH SEXES ADMISSIONS Number	BOTH SEXES ADMISSIONS %
Law	120	96	**80**	80	64	**80**	200	160	**80**
Business	120	24	**20**	80	16	**20**	200	40	**20**
Both schools	240	120	**50**	160	80	**50**			



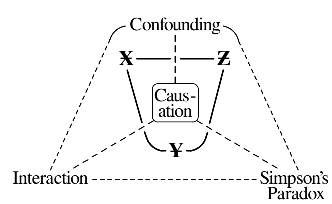
3. Across Tables HL69.1 to HL69.6 on pages HL69.1 and HL69.2 and Tables HL69.13 to HL69.15 overleaf on page HL69.3 and above, different weights in the proportion calculations (like those in Table HL69.7 on page HL69.2) yield a noteworthy *variety* in the percentages for women compared to those for men. This is summarized in Table HL69.16 at the right below; three categories are distinguished.
   - In four tables, there *is* an $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationship, there is no interaction of $\mathbf{X}_1$ and $\mathbf{X}_2$ in their effects on $\mathbf{Y}$ and, in two of the tables, the $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationship is *un*exceptional in light of the effect of subdivision by $\mathbf{X}_2$; by contrast, in Table HL69.3 and Table HL69.6 between the first and second levels of subdivision by $\mathbf{X}_2$, the exceptional behaviour is the $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationship *disappearing* when the data are subdivided by $\mathbf{X}_2$.
     – Notation like (1,3) [or (1,2)] on Table HL69.5 (or Table HL69.6) in Table HL69.16 refers to the first and third (or first and second) levels of subdivision by $\mathbf{X}_2$.
   - In four tables, there is *no* $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationship but three of these are '*false* negative' Answers – when the data are subdivided by $\mathbf{X}_2$, there *is* an $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationship and so they are designated 'exceptional' in the fourth column.
     – In Table HL69.14, interaction is the *reason* for the the exceptional behaviour but interaction is absent in the other three tables.
   - In five tables, there *is* (again) an $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationship, interaction is absent or incidental, and each is a case of the exceptional behaviour known as Simpson's Paradox.

Table HL69.16: $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ **RELATIONSHIPS IN NINE TABLES**
(SP in the fourth column denotes 'Simpson's Paradox')

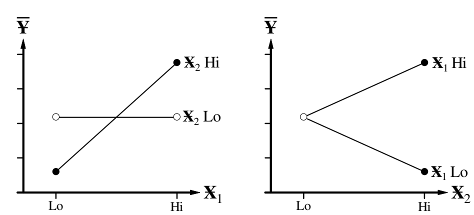
Table	Relationship	Interaction	Exceptional behaviour
HL69.3	Yes	No	Yes
HL69.4	Yes	No	No
HL69.5 (1,3)	Yes	No	No
HL69.6 (1,2)	Yes	No	Yes
HL69.2	No	No	Yes
HL69.6 (2,3)	No	No	Yes
HL69.14	No	Essential	Yes
HL69.15	No	No	No
HL69.1	Yes	No	Yes: SP
HL69.5 (1,2)	Yes	No	Yes: SP
HL69.5 (2,3)	Yes	No	Yes: SP
HL69.6 (1,3)	Yes	No	Yes: SP
HL69.13	Yes	Incidental	Yes: SP

Apart from understanding the properties of weighted averages, the summary in Table HL69.16 above reminds us that:
   - ⊙ Simpson's Paradox is merely the *most* exceptional case (*change of direction* of an $\mathbf{X}_1$-$\bar{\mathbf{Y}}$ relationship) in a context (involving *discrete* variates) that can give rise to less exceptional or even *un*exceptional behaviour;
   - ⊙ interaction is rarely the reason for the exceptional behaviour (only in Table HL69.14 at the top of this page).

4. In meeting the obligation to deal with *relationships* in introductory statistics, the lengthy discussion (*e.g.*, in Statistical Highlights #9, #10, #57 to #70) shows the unexpected) complexities, for only *three* variates, arising from issues of causation, confounding, interaction and Simpson's Paradox. The schema at the right reminds us there are common themes *and* differences among these four matters – see also Appendix 2 on pages HL3.3 and HL3.4 in Statistical Highlight #3.



5. The (equivalent) diagrams at the right show the effects of two (binary) focal variates $\mathbf{X}_1$ and $\mathbf{X}_2$ on (the average of) $\mathbf{Y}$; one focal variate is on the horizontal axis of a diagram, the other distinguishes the two lines by its level.
   - The *non*parallel lines show there is an $\mathbf{X}_1$-$\mathbf{X}_2$ *interaction.*
   - The left-hand diagram shows that $\mathbf{X}_1$ and $\mathbf{Y}$ are *independent* when $\mathbf{X}_2$ is Lo (the relevant line has *zero* slope) but *not* when $\mathbf{X}_2$ is Hi.



   - The right-hand diagram shows there is *no* independence of $\mathbf{X}_2$ and $\mathbf{Y}$ – *neither* line has zero slope.

   For these two dagrams, it is *wrong* to involve *conditional (probabilistic) independence* because it implies that we have undertaken the difficult task of formulating a probability model for this situation – see also Section 13 on page HL89.18 in Statistical Highlight #89.