DOCUMENT RESUME

ED 454 429 CE 081 948

AUTHOR Hawke, Geof
TITLE Explorations of Retention and Withdrawal in VET Programs, 1995-1997. Working Paper.
INSTITUTION Technology Univ.-Sydney, Broadway (Australia). Research Centre for Vocational Education and Training.
SPONS AGENCY Australian National Training Authority, Melbourne.
REPORT NO RCVET-WP-99-21
PUB DATE 1999-00-00
NOTE 26p.
AVAILABLE FROM For full text: http://www.uts.edu.au/fac/edu/rcvet/working%20papers/9921Hawk.pdf
PUB TYPE Numerical/Quantitative Data (110) -- Reports - Research (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Persistence; Developed Nations; Discipline; Dropout Research; Educational Research; Enrollment Trends; Foreign Countries; *Geographic Regions; Majors (Students); *Multivariate Analysis; Regional Characteristics; *School Holding Power; *Sex Differences; Student Attrition; *Vocational Education; *Withdrawal (Education)
IDENTIFIERS *Australia

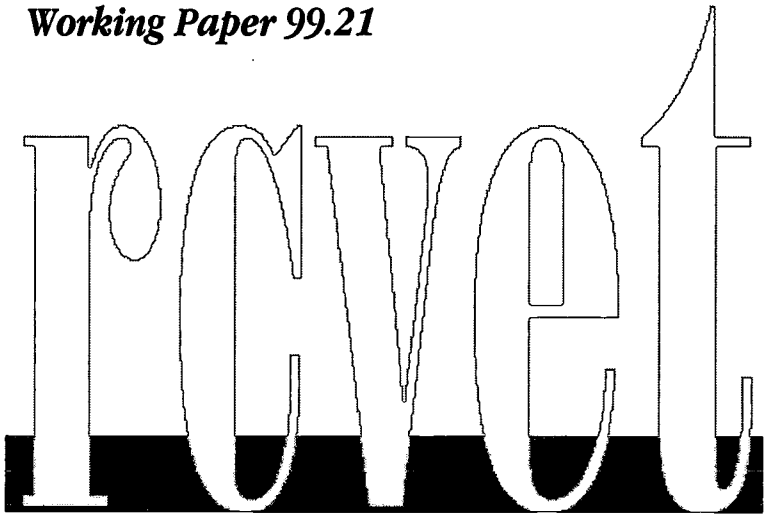
ABSTRACT

An exploratory study was the first part of an extensive study of retention in vocational education and training (VET) modules in Australia. These four major categorical variables were used to explore patterns in the withdrawal rate: year, gender, state/territory, and area of study/discipline. Withdrawal rates declined significantly from 1995-97. The rate at which males and females failed to complete modules differed by slightly less than 1 percent. Differences among states and territories were quite marked--from a withdrawal rate of just over 4 percent in South Australia to more than 21 percent in the Australian Capital Territory. Area of study also greatly affected withdrawal. Differences varied from a low 8.3 percent in Built Environment to a high 18.1 percent for Humanities. Data showed a significant inverse relationship between withdrawal rates for males and females across disciplines. Females had withdrawal rates higher than males in states where withdrawal rates were relatively low; this reversed in states where withdrawal rates are relatively high. A range of methods were explored to understand the interactions among these four variables: hierarchical breakdown analyses; additive analyses (analysis of variance, logit analysis, iterative fits); and weighted net percentage difference analyses. It was found that gender played a role in withdrawal rates only through its mediation by other variables, and that discipline had an important role. (Contains 16 references, 14 tables, and 9 figures.) (YLB)

Explorations of retention and withdrawal in VET programs 1995-1997

Geof Hawke
Senior Research Fellow

Working Paper 99.21



UTS RESEARCH CENTRE FOR VOCATIONAL EDUCATION AND TRAINING
A national key centre supported by the Australian National Training Authority

CONTENTS

Contents 2

Introduction 2

- Purpose of this exploratory study 3
- Limitations of the data 3
- Definitions used in this study 3

The major variables 5

- Year 5
- Gender 5
- State/Territory 5
- Area of Study - Discipline 6

Exploring the data 7

- Hierarchical Breakdown Analyses 9
- Additive Analyses 13
- Weighted Net Percentage Difference Analyses 17

Discussion 19

- Possible data artefacts 19
- Patterns of interest 20
- Future research 22

Bibliography 23

INTRODUCTION

When learners enter a formal course of vocational education, they typically do so in high expectations of achieving the outcome of successful completion. However, we know that a wide range of personal, social, cultural and institutional factors can prevent successful completion.

Over the years in Australia and overseas numerous studies have sought to explore those factors in the hope that "wastage" could be reduced (see, eg Duball & Barrie, 1990, Dunn, 1995 and Mill, 1991). It would be fair to say, however, that most of those studies have focused on factors associated with the individual as if better selection of learners, or more informed choices by learners, could ensure that the learner body was comprised of those "more likely to succeed".

Moreover, most of these studies have considered the factors as distinct variables rather than considering their interactions. This is despite the frequent observation that the factors are highly inter-related.

It would appear that this strategy has not been successful. Concern continues to be expressed that "wastage rates" are too high and that substantial public resources are being "wasted".

It is not our purpose here to debate whether or not any experience of learning should be regarded as "wastage". Rather we seek to explore some of those institutional factors that are more amenable to government intervention.

Purpose of this exploratory study

This study is the first part of a more extensive study that will examine retention in VET modules using a range of techniques. Specifically, the study will:

- Use the publicly available data provided by NCVER through its statistical publications series
- Investigate the patterns of retention using exploratory quantitative approaches that seek to reveal the underlying structure of the data
- Consider the data as inherently complex and involving multivariate interactions
- Explore a range of analytic and exploratory tools to identify those most useful for this data.

Limitations of the data

The published data is inherently constrained by the fact that the complete data is not reported but, rather, that summary tables only are provided. This constrains the number and kinds of interactions that can be explored. In this study, for example, the data is restricted to that involving only four variables:

- The year in which the client enrolled in the module.
- The clients gender
- The coded discipline area in which the module is classified
- The state or territory in which the module enrolment is registered.

Moreover, the data includes only the years 1995-1997 because, though the similar 1998 publication was available, the method of reporting in 1998 had changed in a manner that precluded differentiating by Discipline area studied. Data for 1997 was included to provide a greater diversity over years. However, the precision of 1997 data is significantly less than that for 1995 and 1996 as the '97 data was reported only in multiples of 1,000 module enrolments.

A further issue also needs to be identified. While the definitions used by the states and territories to report client data to NCVER remained essentially constant over that period, there is reason to believe that the interpretation of those classifications may have changed. This is especially believed to have importance in the interpretation of what constituted a "Withdrawal" versus what constituted a "Failure". This may have particular importance for the interpretation of the 1995 data.

Definitions used in this study

This study is focused on retention. However such as outcome is not directly reported in the NCVER data. Rather, eleven possible outcomes are defined. Of these, some are specifically relevant to attendance in modules that are not formally assessed, others relate to modules with formal assessment and others could apply in either case.

To reduce this complexity to manageable levels, some simplifying categorisations have been created:

- a) Pass: 01 - Student assessed-passed
- b) Fail 02-Student assessed-failed
- c) Withdrawn 10-Wiuthdrew-without failure, 11 Withdrew-failed
- d) Withheld 03-Student assessed-result withheld

The withdrawal rate used in this study, then, is defined to be c) divided by the sum of a) to d), above expressed as a percentage. A retention rate could then be defined to be 100-withdrawal rate. However, for convenience it is the withdrawal rate that is used in this report.

Clearly, other means of defining retention and withdrawal could have been followed. For example, other outcomes include codes 04 and 05 in which no assessment applies. As codes 10 and 11 also may apply where no assessment exists an alternative would have been to include these in the divisor. Given the exploratory nature of this study, however, it was more convenient to exclude these at this stage.

The AVETMISS system that provides the data for this study also defines a variable called 'Discipline Group' and this is examined in this study. We use abbreviated titles for the various categories of this variable but a complete listing is provided in the appendix.

THE MAJOR VARIABLES

As noted above, four major categorical variables are used in this study to explore patterns in the *withdrawal rate*. Each of these are ones that are readily available in the published data and for which some reasonable case could be made that ought be significant influences on *withdrawal rates* for various reasons.

In this section, the simple effects of those factors are first described.

Year

A clearly evident pattern is that withdrawal rates declined significantly over the three year period under examination. Figure 1 shows that decline which was substantial from 1995 to 1996 but then with a smaller decline to 1997.

In 1998, it appears that the withdrawal rate has declined further to less than 10.8%. The uncertainty arises because the published data includes a single category 'Withdrawn" that includes the two codes we have used but adds code 12 - Withdrawn- transferred.

Gender

The rate at which Males and Females fail to complete modules in which they enrol differs by slightly less than 1%. Women (14.4%) have a slightly higher rate than males (13.6%).

State/Territory

Differences among the states and territories are quite marked. Over the three years of study, the percentage withdrawing varied from a rate of just over 4% in South Australia to more than 21% in the ACT.

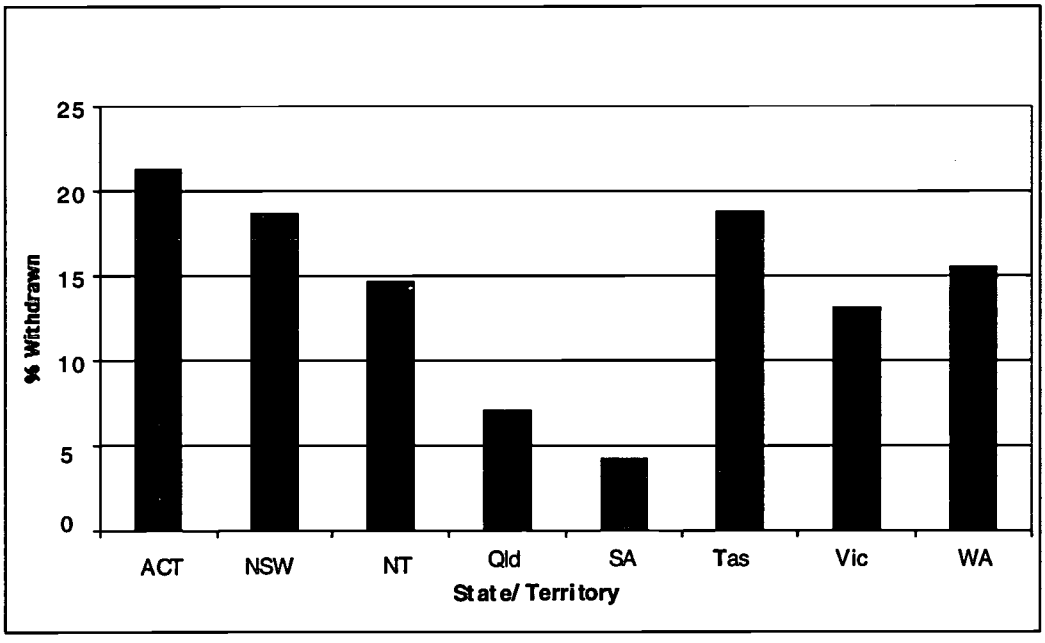


Figure 1 Variations by State/Territory

Area of Study - Discipline

As with the variation between the states, the area of study also greatly affects the extent of withdrawal. Differences here vary from a low 8.3% for those studying in the area of the *Built Environment* to a high 18.1% for the *Humanities*. Figure 3 illustrates the extent of these differences.

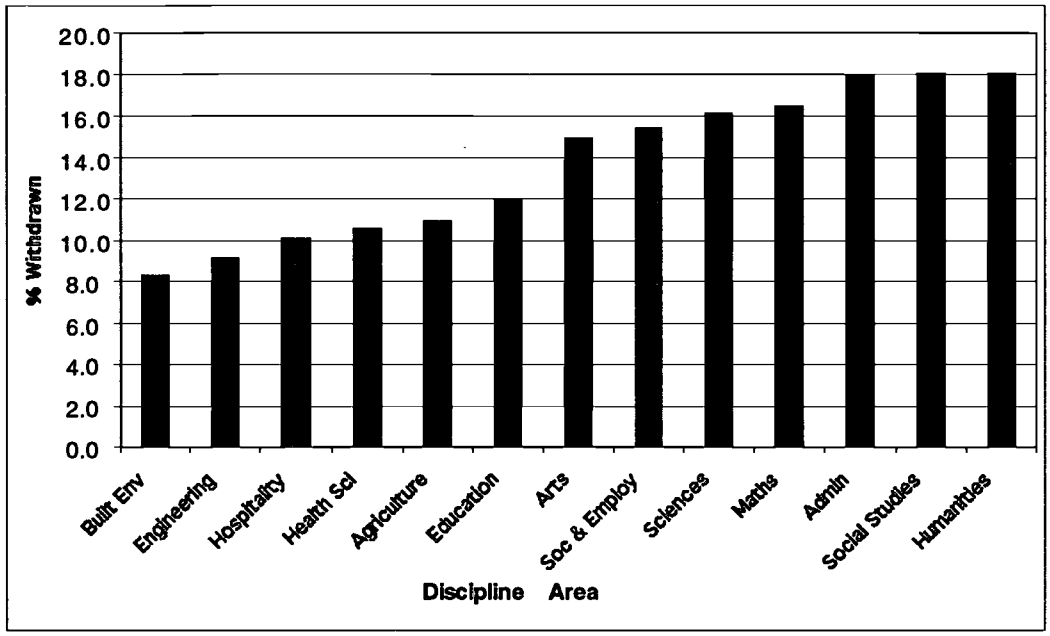


Figure 3 Withdrawal rates by Discipline

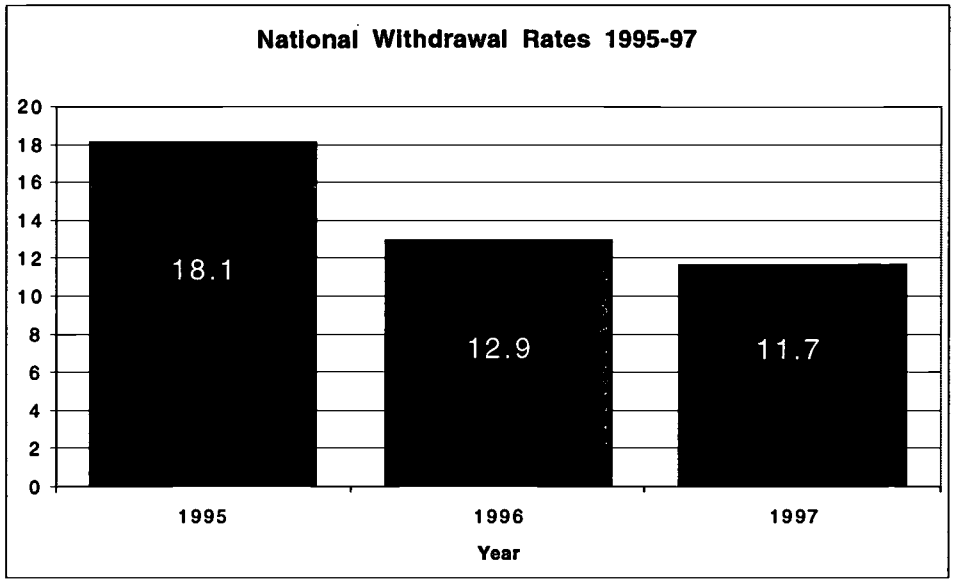


Figure 2 Changes in the rate of withdrawal over years

These gross effects would suggest that while Year, Discipline and State variations might be worth more extensive exploration, Gender does not offer the same degree of interest. However, this would be an erroneous conclusion, for Gender is known to be confounded with Discipline (women are under-represented in certain areas and over-represented in others) but also, it turns out, importantly confounded with State. In both cases, this is most notable at the extremes.

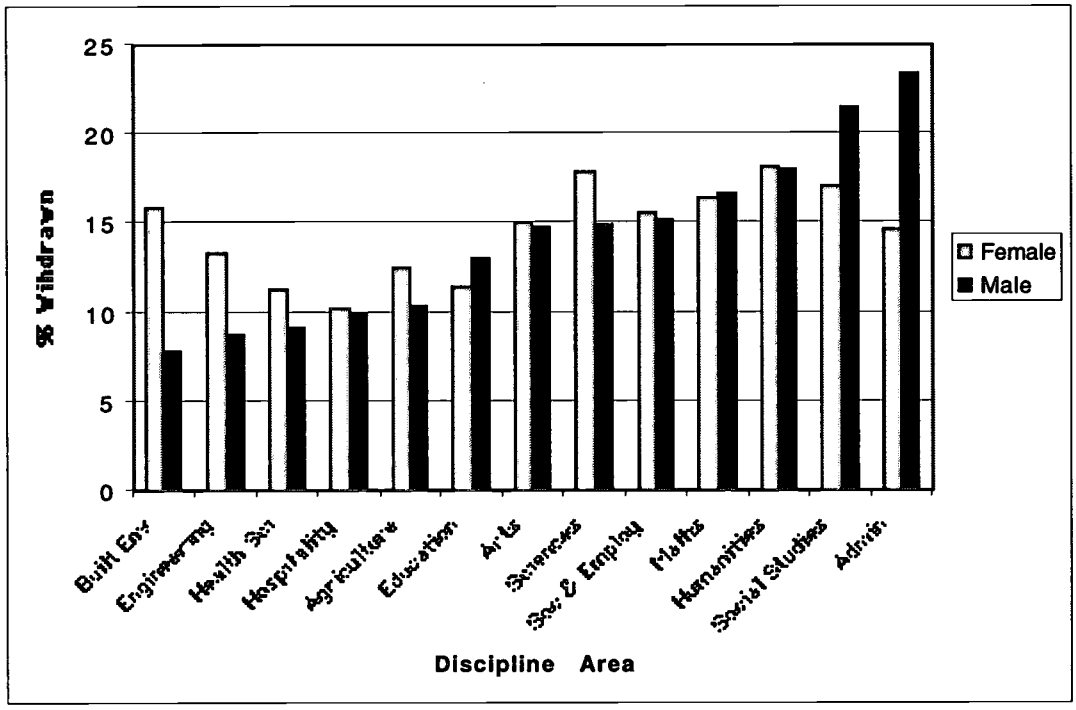


Figure 4 Gender interaction with Discipline

Figure 4 shows that there is a significant inverse relationship between the withdrawal rates for males and females across Disciplines. Figure 5 shows a related pattern in which Females exhibit withdrawal rates higher than their male colleagues in States where withdrawal rates are relatively low but this reverses in those states where rates are relatively high.

Moreover, this pattern has varied over time. Figure 6 shows how male/female rates have varied across the states over the three years.

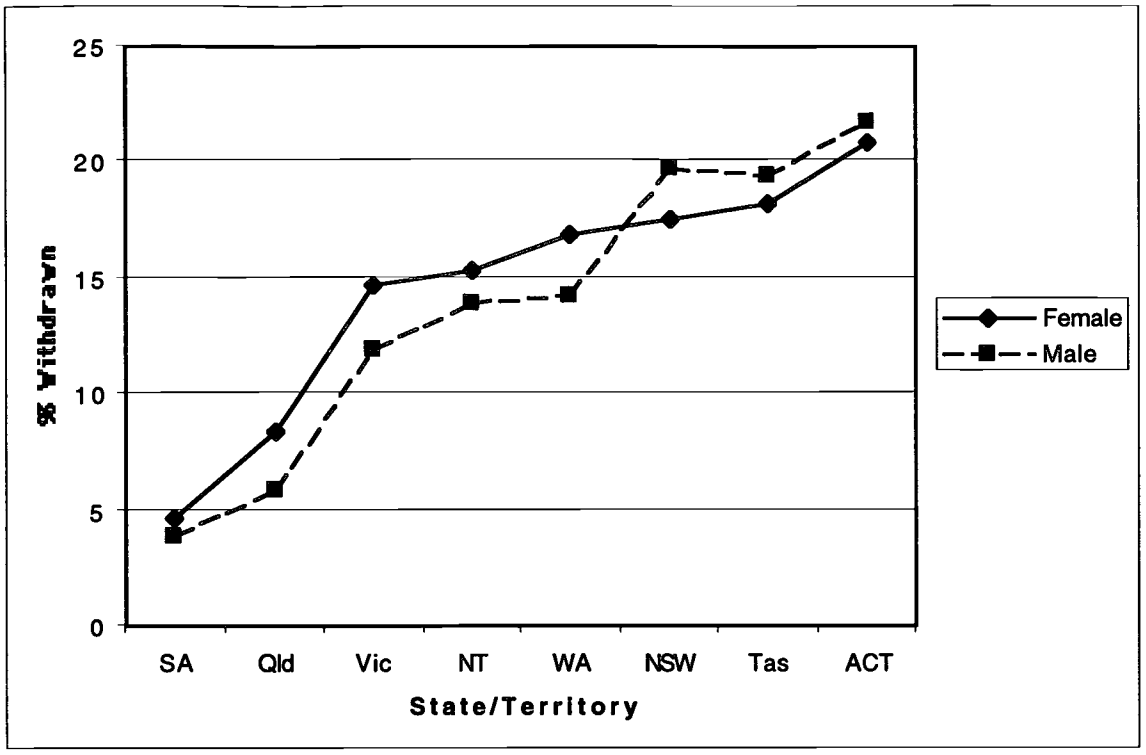


Figure 5 Gender/State interaction

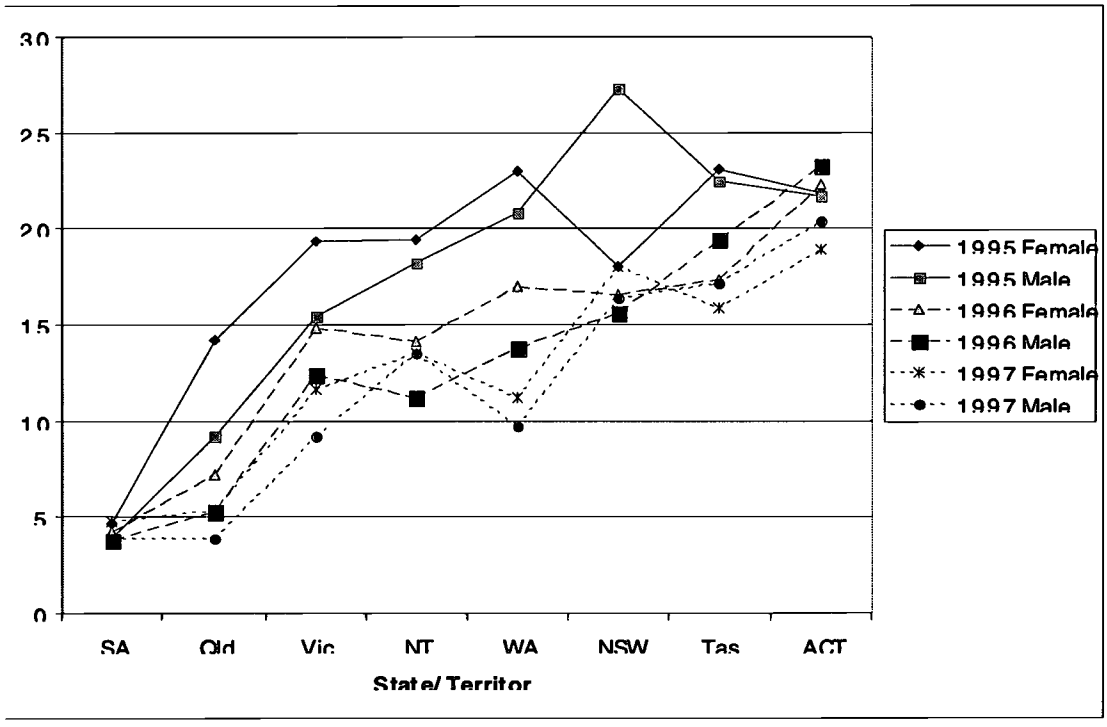


Figure 5 Gender variations over time

The evidence above suggests that the interactions between the four variables are complex. So, in this and later sections, we explore a range of methods for trying to understand these interactions better.

The technique used here is AID one of a class of methods often described as Empirical Decision Trees that are now being used as tools for "data mining" (for other uses in education see Hawke, 1980, Bourke and Keeves, 1977). Essentially the technique involves identifying the division of the current sub-group that leads to a maximal difference in the values of the criterion variable.

In this first examination (Figure 7), we include all four of the predictor variables - *Year*, *State*, *Gender* and *Discipline*. In this case, the criterion variable is the withdrawal rate and this analysis shows that for all module enrolments in the three years, 13.9% of them produced a result of 'withdraw'. However, by looking at the sub-groups that emerge from this analysis a more complex pattern becomes clear.

The first split is on the basis of *State*. Clients enrolling in Qld or SA are seen to have significantly lower withdrawal rates and three distinct groups can be seen. The question that must be asked is whether this is related to the structure or delivery of VET programs in those states or is it in some way an artefact of the data collection. We will return to this later in the discussion section.

For two of these groups of states, the next most significant differentiating variable is *Year*. In both cases, the withdrawal rates in 1995 are substantially greater than those in 1996 or 1997. Again, the question of artefact or real phenomena arises.

For the ACT and Tasmania, however, the Discipline area studied is the next significant factor. In those states/territories, the disciplines divide into two approximately equal sized groups whose withdrawal rates differ by more than 8%.

The importance of *Discipline* as a differentiating factor can also be seen as it appears again in the two groups on the left of the figure. However, of the eleven subgroups in which *Discipline* is a differentiating factor, no two use the same groupings of discipline areas. The only significant pattern to be repeated is that, wherever they appear, the disciplines of Humanities, Maths, Science, Social and Employment Skills and Social Studies cluster together. This suggests that there are similarities in the programs themselves that may influence the rates of withdrawal. However, this is clearly no simple effect as the withdrawal rates for the sub-groups in which they are operative vary from a low 6.1% to a high 26.4%.

It is notable, though, that these disciplines are typically involved in the group, at each splitting point, that has the higher withdrawal rate. That is while their absolute effect varies, their relative effect is towards higher rates of withdrawal.

Another way of summarising the results of this analysis is to look at the final set of sub-groups and Table 1 provides a listing of those.

Hierarchical Breakdown Analysis
Withdrawal Rates 1995-1997

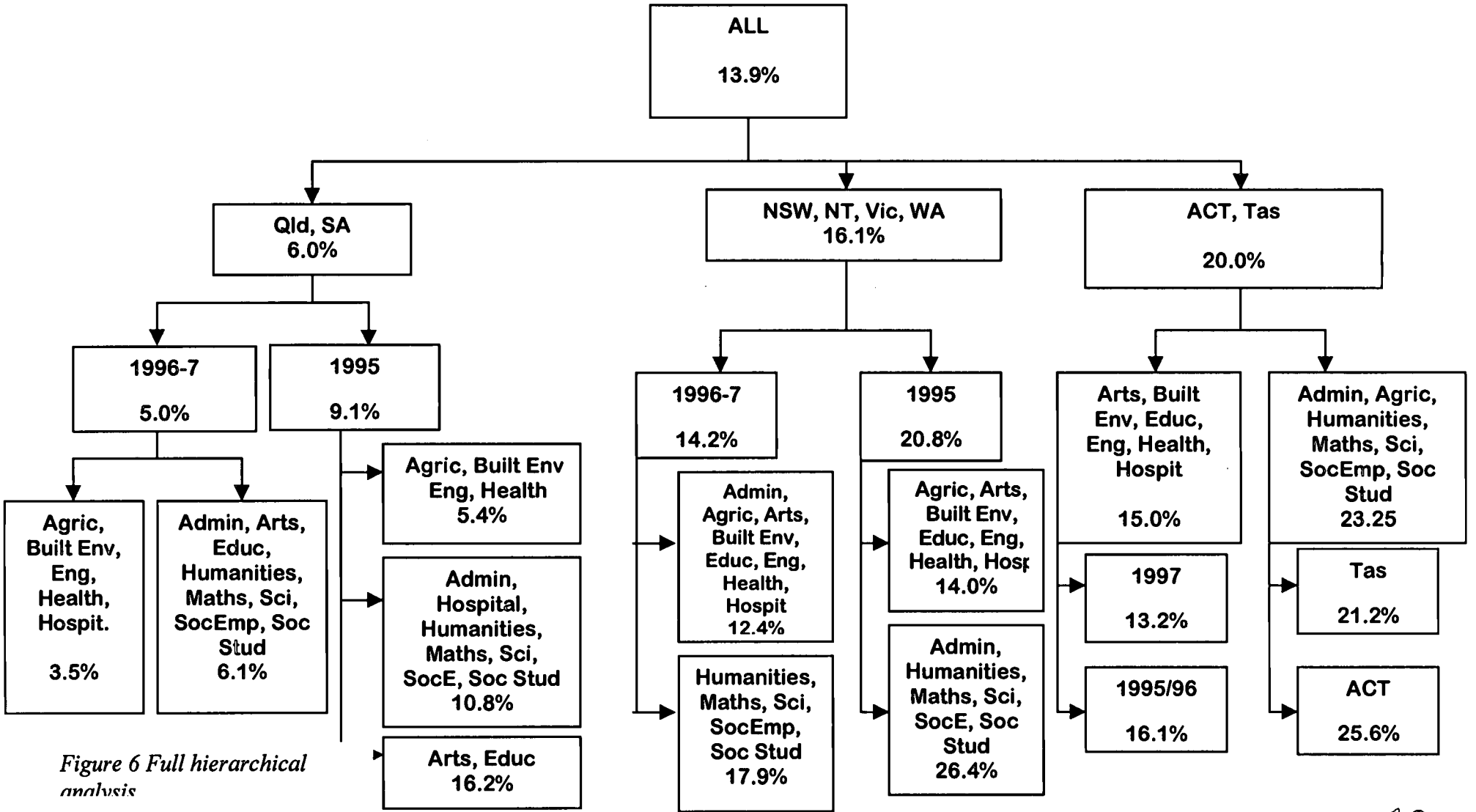


Figure 6 Full hierarchical analysis

Table 1 Sub-groups resulting from breakdown analysis

Sub-Group Description	% Withdrawn
▪ Qld and SA clients in 1995 studying Agriculture, Built Environment, Engineering or Health Sciences	5.4%
▪ Qld and SA clients in 1996 or 1997 studying Administration, Arts, Education, Humanities, Maths, Sciences, Soc & Employment or Soc Studies	6.1%
▪ Qld and SA clients in 1996 or 1997 studying Agriculture, Built Environment, Engineering, Health Sciences or Hospitality	3.5
▪ Qld and SA clients in 1995 studying Admin, Hospitality, Humanities, Maths, Sciences, Soc & Employment or Soc Studies	10.8
▪ NSW, NT, Vic or WA clients in 1996 or 1997 studying Admin, Agric, Arts, Built Env, Health Sci, or Hospitality	12.4
▪ ACT or Tas clients in 1997 studying Arts, Built Env, Education, Engineering, Health or Hospitality	13.2
▪ NSW, NT, Vic or WA clients in 1995 studying Agric, Arts, Built Env, Engineering, Health Sci or Hospitality	14.0
▪ ACT or Tas clients in 1995 or 1996 studying Arts, Built Env, Education, Engineering, Health or Hospitality	16.1
▪ Qld and SA clients in 1995 studying Arts or Education	16.2
▪ NSW, NT, Vic or WA clients in 1996 or 1997 studying Humanities, Maths, Sciences, Soc & Employment or Soc Studies	17.9
▪ Tas clients studying Admin, Agric, Humanities, Maths, Sciences, Soc & Employment or Soc Studies	21.2
▪ ACT clients studying Admin, Agric, Humanities, Maths, Sciences, Soc & Employment or Soc Studies	25.6
▪ NSW, NT, Vic or WA clients in 1995 studying Admin, Humanities, Maths, Sciences, Soc & Employment or Soc Studies	26.4

It is notable that *Gender* does not appear as a differentiating factor in the figure. Indeed within each of the final sets of sub-groups, the differences in withdrawal rates between males and females rarely exceeds 2% (roughly the overall difference). To explore this further, a reduced analysis was conducted using only *Gender* and *Discipline*. Alternatives using *State* sand *Year* did not lead to such great variation, suggesting that *Gender* is more deeply confounded with them than with *Discipline* — an unexpected outcome. Figure 8 reports the results of that analysis.

When restricted to these two factors only, *Discipline* is the factor that provides the initial and greatest variation. In the first instance, it identifies three sub-groups of *Disciplines* that vary substantially in withdrawal rate. The leftmost of these, has a rate of 10% while the rightmost (Admin, Humanities, Maths, Science and Social Studies) has a rate of 17.5%. Each of these sub-groups is then examined in turn and the optimal split identified. In two cases, that split now involves the variable *Gender* with the resultant sub-groups differing in withdrawal rates from a low of 9.2% to a high of 15.5%.

The process continues until all sub-groups have achieved a degree of stability such that further division does not lead to significant change. When that stage had been reached in the present analysis, it resulted in nine sub-groups whose withdrawal rates varied from 9.2% to 21.9% as shown in Table 2.

Hierarchical breakdown analysis
Withdrawal rates 1995-1997 by gender and discipline area

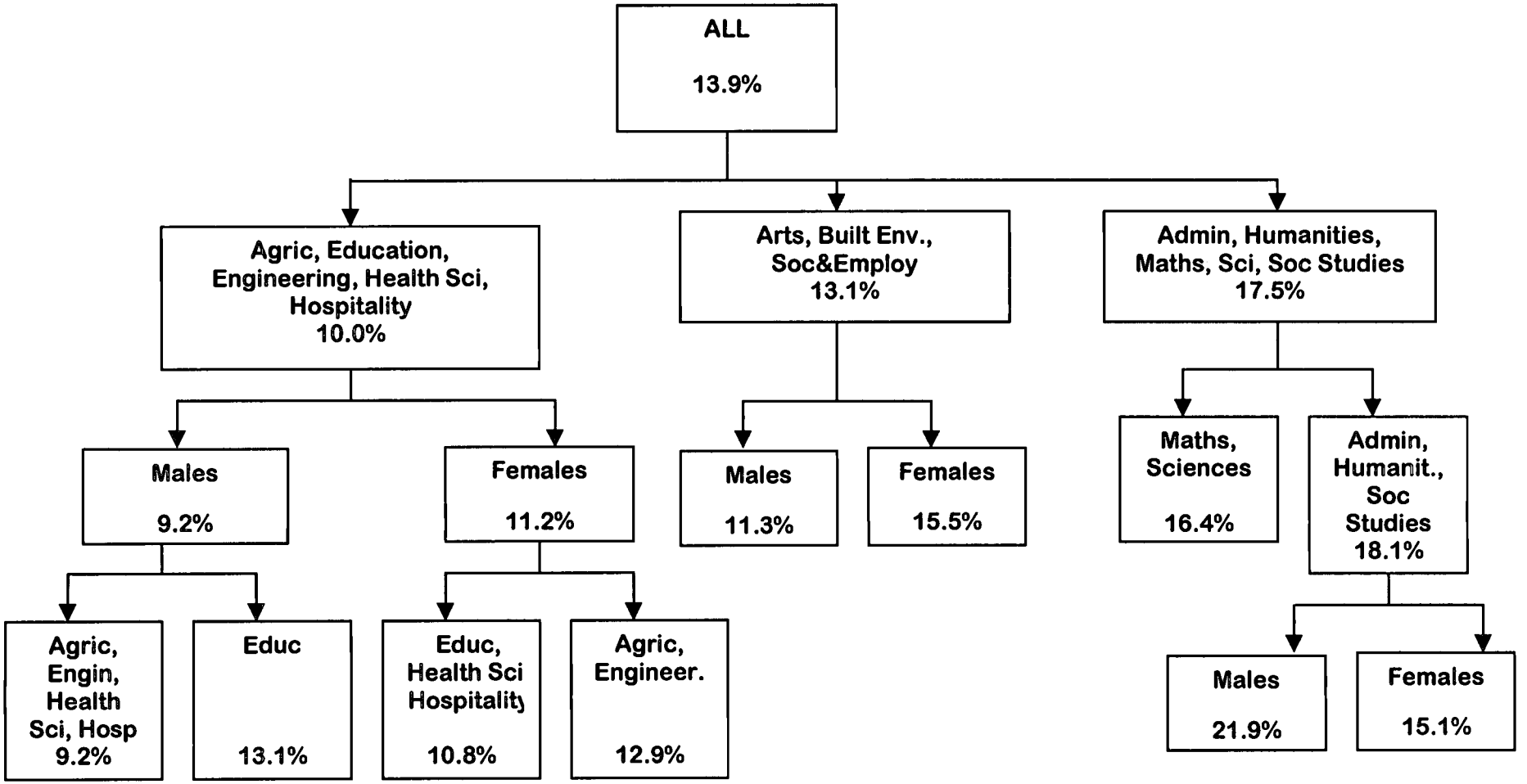


Figure 7 Hierarchical breakdown of withdrawal rate - Discipline & Gender only

Table 2 Gender/Discipline sub-groups

Sub-Group Description	% Withdrawn
Males in Agriculture, Engineering, Health Science or Hospitality	9.2
Females in Education, Health Sciences or Hospitality	10.8
Males in Arts, Built Environment or Social & Employment	11.3
Females in Agriculture or Engineering	12.9
Males in Education	13.1
Females in Admin, Humanities or Social Studies	15.1
Females in Arts, Built Environment or Social & Employment	15.5
All clients in Maths/Computing or Sciences	16.4
Males in Admin, Humanities or Social Studies	21.9

There are a number of observations and questions that arise from this analysis. In particular, what are the characteristics of the three Discipline groups that are the result of the first split that lead to such a diversity of withdrawal rates? Why are Admin, Humanities and Social Sciences the only Discipline areas in which Females show lower withdrawal rates than males?

Additive Analyses

Another way in which this data could be explored is by examining the extent to which the effects of the major variables can be modelled as additive. That is, can we build a model that provides us with an explanation of the contribution of the various factors to the final result. There are many ways in which this can be done but, in this exploratory study, we have focused on one traditional approach and two others that are less well known.

ANOVA

The traditional approach is the Analysis of Variance but, in this case, it will be carried out using a weighted least-squares approach to take into account the very wide differences in the module enrolments in the different Discipline areas as well as State, Gender and Year variations.

The ANOVA table (Table 3) shows that of the main effects and two-way interactions tested, only Gender is not significant. However, Gender is involved in each of the interactions with the other variables. This lends support to the earlier observation that Gender alone does not significantly associate with the withdrawal rate but that it is a significant mediating factor for other variables.

While the table suggests that significant effects can be identified, examination of the model's fit suggests that the additive structure may be misleading and may be more a function of the substantial size of the data set. This is best illustrated by examining the weighted residuals for the main effects and Table 4 illustrates the model's lack of fit for the variable State. Only for SA are the residuals less than 5%.

Table 3 ANOVA - Main effects and 2-way interactions

	DF	SS	MS	F	P-value
CONSTANT	1	19114	19114	330.89946	0
year	2	1278.4	639.21	11.06620	<0.001
state	7	5866.7	838.1	14.50940	<0.001
disc	12	1924	160.33	2.77567	0.001
gender	1	148.01	148.01	2.56237	0.110
year.state	14	2796	199.72	3.45754	<0.001
year.disc	24	6737.3	280.72	4.85989	<0.001
year.gender	2	1717.8	858.91	14.86965	<0.001
state.disc	84	9454.8	112.56	1.94862	<0.001
state.gender	7	2812.5	401.78	6.95580	<0.001
disc.gender	12	5372.5	447.71	7.75088	<0.001
ERROR	458	26455	57.762		

Table 4 Weighted residuals for State

ACT	NSW	NT	Qld	SA	Tas	Vic	WA
-5.2	-23.4	-21.2	-17.4	0.8	-5.2	-18.1	-22.1

Similar outcomes could be provided for each of the other main effects.

Logit Analysis

However, traditional ANOVA is not the best method for analysing rates such as applies here. An alternative method known as logit analysis provides a more appropriate analytical framework. Using an iteratively reweighted least squares approach (McCullagh & Nelder, 1983) as implemented by the program MacAnova (Oehlert & Bingham, 1998), a linear additive model of main effects and 2-way interactions was fitted. Table 5 reports the 'deviance table' from that analysis.

```
Model used is y=year + state + disc + gender + year.state +
year.disc + year.gender + state.disc + state.gender + disc.gender
NOTE: summaries are sequential
```

Table 5 Logit analysis deviance table

	DF	Deviance	MDev	Prob
CONSTANT	1	1.1304e+07	1.1304e+07	<0.001
year	2	1.1441e+05	57207	<0.001
state	7	4.4388e+05	63412	<0.001
disc	12	2.1786e+05	18155	<0.001
gender	1	6941.8	6941.8	<0.001
year.state	14	29409	2100.6	<0.001
year.disc	24	32039	1335	<0.001
year.gender	2	13706	6853	<0.001
state.disc	84	58212	693	<0.001
state.gender	7	20982	2997.4	<0.001
disc.gender	12	36791	3065.9	<0.001
ERROR	458	1.0769e+05	235.13	<0.001

The deviance table reports the outcome of successively fitting models with additional variables or interactions considered. Thus line 5 (labelled "gender") reports the additional fit of the model with a constant term and each of the main effects (i.e including Year, State,

Discipline and Gender) over the model that does not include Gender. The overall analysis of the table shows that the model fits the data well and that each term adds significant improvements to the fit. Clearly, then, we are dealing with a complex pattern of both main effects and interactions.

Iterative Fits

Another approach derives from the exploratory analysis approach originated by John Tukey (Tukey, 1977, Mosteller & Tukey, 1977). Mosteller and Tukey described the method as PLUS Analysis and it essentially involves repeatedly fitting means, medians or other appropriate summaries to the main effects of two or higher dimensional tables until the fit stabilises.

To illustrate the method it is first applied to the two-way table of Year and State. Table 6 shows the raw data and Table 7 the fit for the unweighted means.

Table 6 % Withdrawal by Year and State

	1995	1996	1997
ACT	21.7	22.8	19.7
NSW	23.2	16.1	17.2
NT	18.8	12.7	13.5
Qld	11.5	6.2	4.6
SA	4.2	4.1	4.3
Tas	22.7	18.5	16.5
Vic	17.1	13.5	10.3
WA	22.0	15.4	10.4

Table 7 PLUS fit for the data of Table 6

	1995	1996	1997	Effect
ACT	-2.9	2.3	0.6	**6.9**
NSW	1.1	-1.8	0.7	**4.4**
NT	0.6	-1.4	0.8	**0.5**
Qld	0.8	-0.3	-0.5	**-7.0**
SA	-2.8	-0.1	2.9	**-10.7**
Tas	0.2	0.2	-0.4	**4.8**
Vic	0.2	0.8	-1.0	**-0.8**
WA	2.9	0.4	-3.2	**1.5**
Effect	**3.3**	**-0.1**	**-2.3**	**14.4**

The row and column 'effects' are shown and represent the marginal contribution of each state or year to the overall effect. That is we've fitted a model such that:

% Withdrawal = Common + Year effect + State Effect

The entries in the body of the table are the residuals, ie they represent the misfit of the model. In this case 15 of the 24 residuals are less than 1% and the greatest is 3.2%. That is a reasonable outcome given the nature of the data. However, it is by looking at the effects and the residuals more closely that a number of patterns occur.

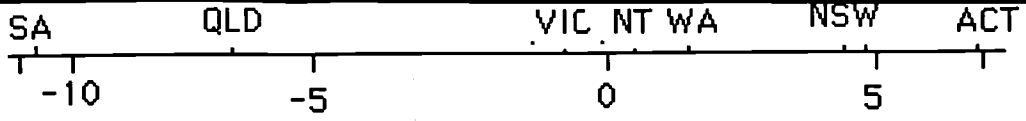


Figure 8 State Effects - PLUS Analysis

Figure 9 shows the relative magnitudes of the State effects. The differences among the states are substantial and this is consistent with the importance of their role in the hierarchical breakdown analysis reported above.

Looking at the residuals, we note that three states account for most of the residuals that exceed 1% - SA, WA and the ACT. For these three states their pattern of changing withdrawal rates over the three years differs from that of the remaining states. In the case of SA, the consistent decline showed by the majority of states does not apply and their rates have remained relatively consistent. The reverse is true for WA while for the ACT, the reported withdrawal rate rose in 1996 before falling in 1997.

We turn now to the same analysis but using the factors of State and Discipline. Table 8 reports on the outcome of this analysis. The effects for State are highly similar to those found in the earlier analysis (Table 7) and, indeed the correlation between the two is 0.996. As well, the estimates of the effects for the factor Discipline are given in Figure 9. As we had noted in the hierarchical analysis, these show a cluster at the positive end of the scale that includes Maths, Sciences, Humanities, Soc and Employment and Social Studies. This is the same group identified there but it is now joined by Admin.

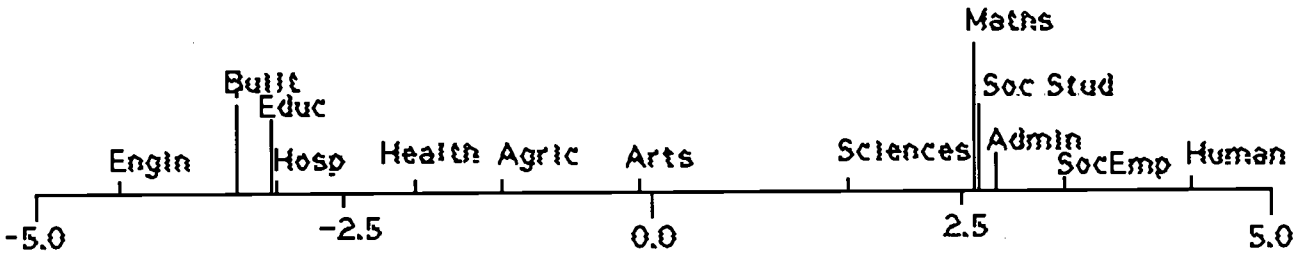


Figure 9 Effect estimates for Discipline

However, in this analysis, the residuals have a much greater range suggesting strongly that the simple additive model is not a good fit and that further fitting should be explored (residuals greater than 4% have been highlighted in the table).

A range of further fits, including adding a product term were explored and diagnostic plots of the residuals were examined. However, it became clear that no improvement added substantially to the simple additive fit.

The conclusion is that the misfitting reflected by the large residuals above reflects real anomalies in the overall pattern of withdrawal rates. Thus the analysis has identified a range of situations in which atypical behaviour has been identified that may warrant further investigation.

Table 8 PLUS Analysis - State and Discipline

	ACT	NSW	NT	Qld	SA	Tas	Vic	WA	Effect
Admin	1.06	4.64	-2.36	-2.05	-2.50	1.12	-0.02	0.11	2.78
Agriculture	-1.36	1.40	-1.31	0.38	-1.38	6.82	-4.63	0.10	-1.22
Arts	-6.35	1.38	0.18	1.74	3.21	-2.76	1.64	0.96	-0.09
Built Env	-2.31	-1.40	4.70	-0.31	2.71	1.39	-4.10	-0.67	-3.36
Education	-1.79	-4.84	-6.31	7.55	-0.49	0.68	2.97	2.22	-3.07
Engineering	-2.88	-2.08	-1.19	0.51	2.17	4.61	1.47	-2.64	-4.32
Health Sci	-1.69	-1.68	3.18	-0.79	1.26	-0.34	-0.45	0.51	-1.92
Hospitality	0.03	-0.70	3.12	3.58	0.88	-5.95	-0.09	-0.87	-3.40
Humanities	4.51	1.00	1.56	-2.20	-3.86	-1.44	-0.55	0.98	4.37
Maths/IT	3.15	-2.58	-1.21	-2.22	-0.62	-1.67	4.07	1.08	2.61
Sciences	2.71	4.09	-2.53	-1.32	-1.40	0.13	-1.04	-0.63	1.58
Soc & Employ	9.46	0.63	-2.04	-1.80	-1.43	-3.22	-2.83	1.23	3.35
Social Stud	-4.54	0.12	4.23	-3.06	1.46	0.61	3.56	-2.39	2.65
Effect	5.20	4.06	0.43	-6.29	-9.4	4.95	-0.75	1.79	13.71

Weighted Net Percentage Difference Analyses

Another simple technique that seeks to identify the effect of the factors was developed by Spady in 1970. It draws on the simple concept of seeking to find the weighted net percentage difference between levels of each factor. Where a factor is binary (as with Gender), the technique leads to a single estimate of the effect of that factor. However, when the factor is nominal (ie categorical as with Discipline) the separate effects of each category are computed. For an ordered variable (eg Year) a single effect estimate is derived that reflects the overall difference assuming linear change.

The importance of the simple algorithm used it that it effectively controls for the effect of confounding variables and thus seeks to fit an additive model with no interaction terms. However, the interim terms of the computation can be used to examine interactions where they exist. An example is the calculation of the effect of Year shown in Table 9. This shows that over the years under study — after controlling for the effect of the other variables — withdrawal rate declined by 5.6%. However, the table also shows the contributions towards that overall effect of various combinations of State, Gender and Discipline. Those that are notably large in absolute terms have been highlighted. The most distinctive pattern to be observed in these relates to the influence of State. Noteworthy here is the status of NSW and SA. NSW can be seen to show significant gender variations across almost all disciplines but the pattern varies — some disciplines show substantial declines for female clients over the three years, but in others substantial increases. For SA, the most notable feature is that there are virtually no differences regardless of discipline or gender.

Similar analyses were conducted for each of the remaining variables and the result of these is summarised in Table 10.

Table 9 WNPD Effect of Year

Effect of Year: WNPD = -5.6%

		Admin	Agric-ulture	Arts	Built Env	Educ-ation	Engin-eering	Health Sci	Hospit ality	Human ities	Maths	Sci-ences	Soc & Emp	Social Studs
ACT	Female	-0.5	-5.4	-6.2	-8.0	3.7	-8.5	-3.3	-2.1	0.4	-2.9	-6.1	-5.0	-1.6
	Male	0.4	-7.2	6.4	-0.5	-26.9	-4.4	-6.7	3.7	2.0	-1.0	-1.8	-6.5	-1.1
NSW	Female	0.5	-55.4	-58.6	19.2	-56.7	-171.3	5.3	-40.5	25.0	80.7	352.9	-49.9	91.4
	Male	-47.8	-1.9	-2.7	0.0	2.9	-0.3	-0.4	-0.1	1.5	1.8	2.6	-0.8	-1.6
NT	Female	-1.6	1.3	-9.3	-14.3	-10.0	-6.8	-15.0	-7.2	-16.3	-0.1	-12.8	-3.5	3.6
	Male	-3.5	-0.3	7.3	-9.0	-29.4	1.6	-17.1	2.0	-15.0	-1.7	-4.9	0.4	20.3
Qld	Female	-8.5	-6.7	-10.8	-7.7	-8.7	-7.9	-5.9	-6.1	-8.9	-9.7	-7.8	-9.4	-6.6
	Male	-5.7	-5.2	-7.3	-2.2	-20.1	-3.3	-4.5	-4.6	-7.1	-7.1	-5.3	-9.5	-11.4
SA	Female	-0.6	-0.4	-0.6	-0.9	0.0	0.4	1.1	-0.1	0.0	-0.5	2.9	1.5	2.4
	Male	-0.2	0.5	-1.1	0.2	-1.7	-1.0	-0.1	0.2	-0.5	-1.0	-1.6	0.4	0.7
Tas	Female	-8.0	-3.3	-3.4	-3.8	-14.8	-3.2	-1.0	-2.4	-12.2	-5.9	-6.0	-9.1	-8.0
	Male	-8.5	-5.3	-2.9	-7.7	-10.5	-1.9	-4.0	0.4	-10.4	-5.3	-14.3	-5.0	-6.6
Vic	Female	-8.0	-7.4	-8.3	-9.3	-4.8	-9.9	-4.8	-4.8	-6.9	-7.7	-9.5	-7.3	-8.3
	Male	-8.5	-3.9	-7.3	-2.3	-3.5	-6.5	-5.2	-1.7	-8.6	-9.9	-7.2	-7.0	-10.3
WA	Female	-12.2	-8.7	-10.0	-5.5	-13.5	-6.2	-10.0	-17.4	-15.0	-13.5	-10.5	-8.2	-7.4
	Male	-15.5	-9.9	-9.9	-7.2	-27.5	-6.7	-7.4	-8.4	-17.3	-13.4	-12.3	-10.3	-8.9

Some notable features of these analyses are that:

- The correlation between the WNPD effects for State in Table 10 and those found in the earlier PLUS analysis is high - 0.993. This provides useful confirmatory evidence.
- The correlation between the WNPD effects for Discipline in the two analyses, similarly, is exceptionally high at 0.963. Again this provides further confirmation.

Table 10 Summary of WNPD analyses

Year														
	-5.6%													
Gender														
	-0.2%													
State														
	ACT	*NSW*	*NT*	*Qld*	*SA*	*Tas*	*Vic*	*WA*						
	6.0%	6.0%	0.6%	-8.1%	-11.2%	5.2%	-0.8%	1.5%						
Discipline														
	Admin	*Agric*	*Arts*	*Built Env*	*Educ*	*Engin*	*Health Scienc.*	*Hospit*	*Human*	*Maths*	*Sci*	*Soc Emp*	*Social Stud*	
	4.2%	-2.2%	0.9%	-2.8%	-2.2%	-4.3%	-3.4%	-3.8%	3.9%	2.4%	2.3%	2.9%	3.9%	

DISCUSSION

The analyses reported above were intended to be exploratory. They sought to identify some of the interesting factors that might warrant deeper exploration and to examine the usefulness of a range of statistical techniques that may assist in exploring the data.

However, they have served to highlight the complex pattern of interrelationships that exist. It is clear that examination of the simple effects of the four variables used in this study would lead to an overly simplistic understanding of the patterns of retention. Rather, differences in withdrawal rates are a result of a complex interplay of factors. Some of these are likely to be artefacts of the data collection system itself, others may reflect genuine systemic phenomena that may be susceptible to intervention.

It would seem to be useful to explore some of these issues more fully before considering the approach to be taken with further analyses of more complete data sets.

Possible data artefacts

The most significant source of contamination of the data appears to be reflected in the effects of *Year* and *State*. For *Year*, we have seen a substantial decline in the withdrawal rate from 1995 to 1996 and a further, but smaller , decline from 1996 to 1997. The pattern of the decline, however, is not consistent between states/territories. For some, the overall decline is great but others show almost no change. Among those with substantial changes, most fall most quickly between '95 and '96, but others have their falls from '96 to '97. Why is this the case? It is unlikely that these reflect changes in the underlying characteristics of the client population. More likely would be explanations based upon improvements in the quality of service provision. However, while such improvements might have contributed to the fall, the magnitude of change makes it unlikely that is the primary explanatory factor.

For that, it is useful to understand that 1995 represents the first year in which all states and territories provided full and comprehensive module outcomes data. Consequently, the data definitions used were new to most of those — teachers and trainers — who were the ultimate source of the information.

With respect to the variable *State*, a similar position appears. The differences between states/territories in their apparent retention rates seems to great to be attributable solely to differences in the characteristics of the clients or characteristics of the state's delivery system. More likely are explanations that, at least in part, consider that the interpretation of the AVETMISS outcome codes has varied substantially between the states and territories.

There is some evidence that this is the case and that it has taken some years for those with primary responsibility for data provision to become familiar with, and consistent in the use of, module outcome classifications. For example, a Victorian study in 1998 (Cleary & Nicholls, 1998) found that "Institutes are unclear about how completions are being defined" (p.20) and 'that the information about completions generated by the current national system ... is unreliable and inaccurate" (p.20). That study found that identical outcomes were coded quite differently across and within Institutes and that particular codes were used to record different outcomes across Institutes.

The importance of this is that, even by 1998, it is possible that significant inconsistencies in AVETMISS codings will confound or hide genuine systemic factors that influence retention patterns. Thus, use of analyses such is this must be regarded as indicative rather than

conclusive — for example, much of the structure of Fig. 7 may represent these artefacts, leaving only the *Discipline* factor as significant. The data can, however, highlight patterns that warrant further exploration.

Patterns of interest

Patterns of gender

The consistent pattern in which *Gender* is found to play a role only through its mediation by other variables is one of the most interesting outcomes of this exploratory study. The question of gender equity has been a major one for VET systems throughout the last decade and this study suggests that, in general when other factors are held constant, there are few notable differences between male and female withdrawal rates. However, it is important to also consider the areas in which that generalisation breaks down.

When the effects of *Year* and *State* are controlled, then *Gender* does become a factor within *Discipline* sub-groups. To highlight this, the data of Figure 8 is presented in a different format in Table 11.

Table 11 Gender differences within Discipline groups

Discipline(s):	Agriculture, Education, Engineering, Health Sciences, Hospitality	V&P Arts, Built Environment, Social & Employment Skills	Maths/IT, Sciences	Administration, Humanities, Social Studies
Female	11.2%	15.5%	16.6%	15.1%
Male	9.2%	11.3%	16.3%	21.9%

Remembering that the raw *Gender* difference is only approximately 1%, only in the second and fourth of these sub-groups is there a difference worthy of further exploration. Moreover, it is noteworthy that in these two groups, the direction of the difference is reversed. The questions that emerge are of two kinds. Firstly, what are the characteristics of the disciplines in each group that are associated with gender differences of this magnitude? But secondly, are these groups homogeneous in those characteristics? Group four, for example, pairs Administration with Humanities and Social Studies — on the face of it, an unlikely combination. These differences need to be examined in more detail if they are to be understood in ways that can be useful.

Patterns of state differences

While, as we've noted, we may need to be cautious in our interpretation of effects involving *State*, some consistencies are worth mentioning. Firstly, a number of analyses lead to roughly similar clusterings and these have an interesting interaction with *Gender* as Table 12 shows.

Table 12 State/Territory clusters

Full Hierarchical Analysis (Fig. 7)	State Effect Estimates - PLUS Analysis (Fig. 9)	Gender-State Interaction (Fig. 5)
◆ Qld, SA	◆ Qld, SA	◆ Qld, SA - low withdrawal, females>males
◆ NSW, NT, Vic, WA	◆ NT, Vic, WA	◆ NT, Vic, WA - medium withdrawal, females>males
◆ ACT, Tas	◆ ACT, NSW, Tas	◆ ACT, NSW, Tas - high withdrawal - males>females

Given that there appears to be more to this grouping than artefact alone, what is it about these groups that is associated with these different patterns of withdrawal? In particular, what is it that explains why the three high rate states are unique in experiencing male withdrawal rates higher than those of females? An obvious line of investigation would be to look at the balance of females among all clients in each state. This shows some relationship but by no means a perfect one:

- Qld and SA have low ratios of female : male clients - both 0.89.
- NT & WA have the highest ratios - 1.04 and 1.07, but Vic has the lowest - 0.79.
- ACT, NSW and Tas have intermediate figures - 0.93, 0.96 and 0.88.

Nor is any discernible pattern evident in the offerings of the states across the 13 discipline areas.

Patterns involving the disciplines

All of the analyses have shown that *Discipline* has an important role in the pattern of withdrawal rates. However, the different analyses have highlighted somewhat differing patterns as Table 13, which lists the dominant clusterings of disciplines in each of three different analyses, shows. There is nonetheless a degree of consistency in linking together Administration, Humanities, Maths/IT, Sciences and Social Studies the five Discipline areas with the highest raw withdrawal rates (see Fig. 3). Frequently these combine with Social and Employment Skills, with the next highest. This tendency to exhibit high withdrawal rates, then, appears to be one that operates regardless of any other factor. Is there, then, a consistent set of factors that underpin this or do unique factors operate for all or some of them?

Table 13 Comparison of Discipline Clusters

Full Hierarchical Analysis (Fig. 7)	Hierarchical breakdown - Gender & Discipline only (Fig. 8)	Effect Estimates from PLUS Analysis (Fig. 10)
▪ Administration, Humanities, Maths/IT, Sciences, Soc & Employment Skills, Social Studies	◆ Administration, Humanities, Maths/IT, Sciences, Social Studies	◆ Administration, Humanities, Maths/IT, Sciences, Soc & Employment Skills, Social Studies
▪ Agriculture, Built Environment, Engineering, Health Sciences, Hospitality	◆ Agriculture, Education, Engineering, Health Sciences, Hospitality	◆ No other standout cluster
▪ V&P Arts, Education	◆ V&P Arts, Built Environment, Soc & Employment Skills	

Probing the areas of poor fit

Several of the analyses reported earlier have attempted to find a model that fits the data. Generally, these also produce examples of particular situations where the model does not fit so well. Table 8, for example, in which we reported the results of the PLUS analysis for the variables *State* and *Discipline* identified some seventeen combinations of the variables where a degree of mis-fit was suggested. These are circumstances where some deeper exploration may give further indications of what underpins the particular pattern being observed.

To illustrate these deeper patterns, we have looked more closely at the four mis-fits in the first column of that table. These represent particular discipline areas in the ACT that are not adequately represented by the simple model of this form of analysis.

Table 14 Exploring some mis-fitting cases

Vis/Perf.Arts	95	96	97
Female	13.4	13.7	8.3
Male	11.2	12.5	16.7

Humanities	95	96	97
Female	23.3	26.8	25.0
Male	28.4	32.4	31.4

Soc & Empl.	95	96	97
Female	31.4	35.5	26.4
Male	35.3	34.9	28.6

Soc Studies	95	96	97
Female	14.7	17.3	13.3
Male	26.7	21.2	25.0

In the case of the Visual & performing Arts, we find an interesting pattern in which for 1997, the female client withdrawal rate sharply declines to be matched by a sharp increase in the male withdrawal rate. This suggests that, for this client group, the Year 1997 exhibited some specific characteristics that had an effect on retention that was gender specific but cannot be further detailed with this limited data set.

For Humanities and Social and Employment Skills, the withdrawal rate is uniformly high suggesting that these modules as delivered in this state may have a specific set of characteristics that are associated with high withdrawal. For Social Studies, clear gender differences are operating. It is worth noting here that males enrolled in Social Studies modules comprise only 19% of clients.

FUTURE RESEARCH

This exploratory work has identified a range of issues that warrant further investigation. As such, it represents a useful step towards a better understanding of the factors that underpin retention. Key questions for further research include the following:

- Can the data (or its analysis) be developed to allow us to distinguish coding and interpretation artefacts from underlying phenomena of greater policy interest?
- What are the underlying factors that are visible as differences in retention rates between disciplines?
- Does the inclusion of other variables such as client age and prior educational attainment assist in clarifying the patterns identified here? In particular, does the minimal effect of gender remain when other factors such as these are taken into account?
- What differences in delivery systems or client characteristics are associated with the differences in retention between states and territories?

BIBLIOGRAPHY

Belsen, W 1959 'Matching on the principle of biological classification', *Applied Statistics*,

Bourke, S F and Keeves, J P 1977 *The mastery of literacy and numeracy: Final report.* Australian studies in school performance, Vol. 3, Hawthorn, Vic, Australian Council for Educational Research.

Cleary, M & Nicholls, R 1998 *Understanding and improving module completions*, Melbourne, Office of Training and Further Education, Victoria.

Duball, B & Baker, R 1990 'Attrition in TAFE: The influences of institutional variables', *Australian Journal of TAFE Research and Development*, vol.5, no.2, pp.17-25.

Dunn, K 1995 'Attrition rates/reason for withdrawal from Adult Basic Education classes: A tentative look at attrition rates in Australia and particular situations overseas including a very small sample survey from Western NSW", paper delivered to NSW TAFE Research Association Conference, Sydney Institute of Technology, June.

Hawke, G A 1980 *Techniques for cross-break analysis*. Tech Notes, Assessment R&D Unit, NSW Department of Technical and Further Education.

McCullagh, P & Nelder, J A 1983 *Generalised linear models* Chapman and Hall, London.

Mill, R 1991 *Study of student attrition in a first year electronics course at a Victorian metropolitan college*, Dissertation for Master of Curriculum Studies, University of New England, Armidale.

Mosteller, F & Tukey, J W 1977 *Data analysis and regression: A second course in statistics* Addison-Wesley, Reading MA.

National Centre for Vocational Education Research 1996 *Selected Vocational Education and Training statistics 1995* NCVER, Leabrook, SA.

National Centre for Vocational Education Research 1997 *Australian Vocational Education and Training statistics 1996: In detail,* NCVER, Leabrook, SA.

National Centre for Vocational Education Research 1998 *Australian Vocational Education and Training statistics 1997: In detail,* NCVER, Leabrook, SA.

National Centre for Vocational Education Research 1999 *Australian Vocational Education and Training statistics 1998: In detail,* NCVER, Leabrook, SA.

Oehlert, Gary W & Bingham, Christopher 1998 *MacAnova user's guide.* Technical Report Number 617, School of Statistics, University of Minnesota.

Spady, W G 1970 'Simple techniques for multi-variate analysis', *Interchange*, vol.1, pp.3-20.

Tukey, J W 1977 *Exploratory data analysis* Addison-Wesley, Reading MA.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Explorations of retention & withdrawal in VET programs 1995-1997 (WP 99-21)	
Author(s): Geof Hawke	
Corporate Source:	Publication Date: 1999

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1 ↑

☑

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A ↑

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B ↑

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Knight	Printed Name/Position/Title: Linda Knight
Organization/Address: RCVET, UTS.	Telephone: 0295143700 FAX: 3737
	E-Mail Address: Date: 18/6/01.

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

Acquisitions Coordinator
ERIC Clearinghouse on Adult, Career, and Vocational Education
Center on Education and Training for Employment
1900 Kenny Road
Columbus, OH 43210-1090

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to: