# APPLETS TO SUPPORT REASONING ABOUT EXPLAINED AND UNEXPLAINED VARIABILITY

MORPHETT, Anthony and GUNN, Sharon
School of Mathematics and Statistics
The University of Melbourne
a.morphett@unimelb.edu.au, sharonkg@unimelb.edu.au

*Reasoning about overall variability in a data set, and the partitioning of variability into* explained *and* unexplained, *is a key aspect of statistical reasoning. It also provides a powerful conceptual framework for unifying several sometimes disparate topics of introductory statistics such as linear regression, ANOVA, and significance testing of a fitted model. This type of reasoning creates a platform for comparing statistical models and hence lays a foundation for inferential thinking. We will describe a pair of applets that have been developed with the explicit aim of helping students to reason visually about the partitioning of variability into explained and unexplained variability. The first applet deals with the case of a numeric explanatory variable (regression for bivariate data), the second with the case of a categorical explanatory variable (single-factor ANOVA). We highlight some aspects of the applets' design which are of pedagogical relevance, and discuss how the applets may support possible learning trajectories for reasoning about variability in the topics of regression, ANOVA and statistical modelling.*

## INTRODUCTION

The partitioning of variability into a portion that is *explained* by a statistical model (a.k.a. 'signal', 'between-groups') and a portion that is *unexplained* (a.k.a. 'noise', 'within-groups') is at the heart of statistical modelling. It is the foundation of analysis of variance and inference using the F-test. Nevertheless, it is not often treated explicitly in educational resources. One exception is the work of (Reid & Reading, 2010), (Reid, Reading, & Ellem, 2008), who suggested a framework for describing the progression in thinking and reasoning about explained and unexplained variability and developed associated assessment items to measure the progression.

Given the importance of partitioning variability, and the seeming lack of relevant, accessible resources online, a pair of applets were developed for visualising the partitioning of variability in two contexts: data with a numeric response and numeric explanatory variable (regression), and a numeric response and categorical explanatory (one-way ANOVA). The applets were developed as part of a wider project (Morphett, Gunn, & Maillardet, 2015) and are available at www.melbapplets.ms.unimelb.edu.au . As well as providing visualisations of the partitioning of variability in each case, we wanted visual representations that would reveal the underlying links between the two contexts – in particular, that one-way ANOVA and regression are two instances of the general analysis of variance procedure. The applets' design was informed by research such as the theory of multimedia learning (Mayer & Moreno, 2003). Many of the principles which guided the applets' design were consistent with those discussed in (Wild, Pfannkuch, Regan, & Parsonage, 2013) and (Arnold, Pfannkuch, Wild, Regan, & Budgett, 2011).

The applets share some features with earlier related applets, such as those of the CAST (see http://cast.massey.ac.nz/core/index.html?book=agExper&page=sec_oneFactorAnova , http://cast.massey.ac.nz/core/index.html?book=regn&page=sec_regnAnova) and WISE projects (see http://wise.cgu.edu/portfolio/demo-regression-and-correlation/), but differ in purpose and design. Our applets are intended to support *visual reasoning* (Dreyfus, 1991), as a complement to symbolic reasoning. We believe that visual and symbolic reasoning, married with a narrative, can potentially lead to more complex knowledge structures and consequently, deeper learning. Hence, the applets were designed from the outset to support *conversations*, in lectures or small-groups, and we envision their use accompanied by a spoken narrative. They are not intended to be used in isolation, nor to be a comprehensive treatment of the concept. In the next sections of this paper, we will describe the applets and highlight some aspects of their design which are pedagogically relevant for partitioning of variability.

The applets are designed to support reasoning about explained and unexplained variability, and are most effective when situated within a curriculum that focuses upon statistical modelling with

In H. MacGillivray, M. Martin and B. Phillips (Eds.). Proceedings of the Ninth Australian Conference on Teaching Statistics, December 2016, Canberra, Australia.

models comprised of deterministic and non-deterministic components. In the final section, we describe a possible learning trajectory, making use of the applets, for developing an understanding of modelling in this perspective.

## APPLET 1: VISUALISING PARTITIONING OF VARIABILITY IN ANOVA

The applet initially shows an individual value plot of bivariate data with a categorical explanatory variable (group – on horizontal axis) and numerical response (vertical axis). The data is divided into three groups, shown in red, green and blue, with 5 observations $y_{i1}, \ldots, y_{i5}$ in each group (i = 1, 2, 3). The group mean $\bar{y}_i$ of each group is shown as a triangle aside the group data. (The triangle can be thought of as the pivot of a see-saw, visually re-enforcing that the mean is a 'balance point' for the data.) To the left of the plot's vertical axis is a dotplot (running vertically) of the pooled data. The points in the dotplot are coloured according to group. The grand mean $\bar{\bar{y}}$ is shown as a dashed horizontal line running across the plot and as a black triangle beside the dotplot. To the left of the dotplot is a "variability meter". It is a bar whose height is proportional to the variance $\mathrm{Var}(y_{ij})$ and represents the total variability in the pooled data. It is initially coloured grey. The components just described make up the region labelled 'Overall variability' in Figure 1.

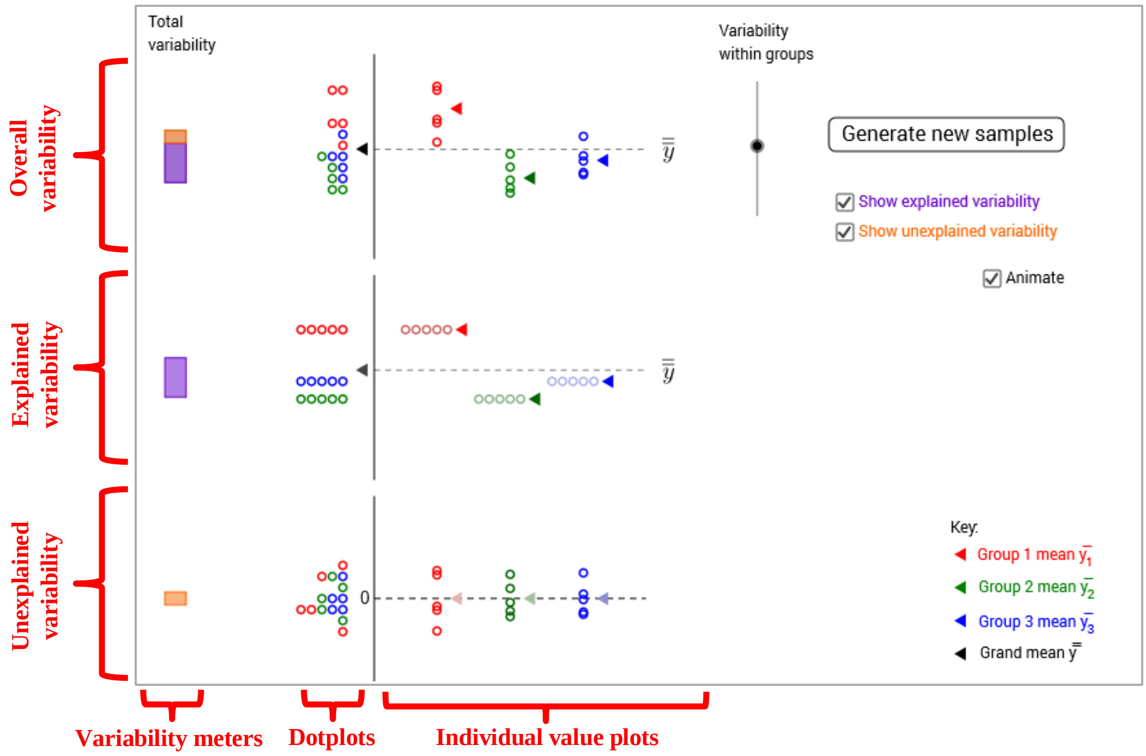



**Figure 1: Screenshot of ANOVA applet after all animations have finished. Sections corresponding to overall, explained and unexplained variability are indicated at the side; the variability meters, individual value plots and scatterplots are indicated beneath.**

The user may modify the group means by clicking and dragging the triangles representing the mean for each group, or may change the within-group variability using a slider labelled 'Variability within groups' to the right of the plot. A button "Generate new samples" produces a new set of data with the same group means and within-groups variance. When the data is modified in any of these ways, the individual value plot, dotplot and variability meter update automatically.

There are two checkboxes to the right of screen. When the first checkbox, labelled 'Show explained variability', is enabled, an animation is triggered. First, a new vertical axis appears in the middle section of the display below the 'Overall variability' section described above. Second, the points in each group of the individual value plot move into a line next to the group mean. Next, the aligned points, group mean triangles, and grand mean line $\bar{\bar{y}}$ descend, maintaining their relative positions, to form a plot by the new axis. The resulting plot shows the group means and 'predicted' values $\hat{y}_{ij} = \bar{y}_i$ of the response, and the grand mean $\bar{\bar{y}}$. Finally, a dotplot of the predicted values $\hat{y}_{ij}$

appears to the left of the new axis, with the points coloured by group; a variability meter for the 'explained' variability appears to the left, and the original data points and means reappear on the top plot. The height of the variability meter is proportional to Var($\hat{y}_{ij}$), or equivalently, to $n$ Var($\bar{y}_i$) where $n$ = 5 is the number of observations in each group. The variability meter is coloured purple, and as it appears, a corresponding portion of the 'overall' variability meter from the top section, equal in height to the new variability meter, is also coloured purple. The animations just described produce the region labelled 'Explained variability' in Figure 1; its structure mirrors that of the overall variability section.

When the second checkbox 'Show unexplained variability' is checked, a set of axes appears in the bottom section of the display, and the data points from each group, and their group means, descend until the group means align with the horizontal axis of the new set of axes, thus forming a plot of the residuals $e_{ij}$ in the ANOVA model $y_{ij} = \bar{y}_i + e_{ij}$. The data points retain their original positions relative to the group mean. To conclude the animation, a dotplot and variability meter appear, and the original data reappears in the top plot. The variability meter in this case is coloured orange, and its height is proportional to Var($e_{ij}$). When it appears, the remaining portion of the overall variability meter from the top section, equal in height to the new variability meter and until now coloured grey, also turns orange. A screenshot of the applet, with all animations complete, is shown in Figure 1.

## APPLET 2: VISUALISING PARTITIONING OF VARIABILITY IN REGRESSION

The applet initially shows a scatterplot of bivariate numerical data $(x_i, y_i)$. The sample size of $n$ = 15 is fixed. The individual data points can be changed by dragging each point within the scatterplot. Also shown on the scatterplot is a horizontal dashed line representing $\bar{y}$, the grand mean of the observed responses $y_i$, and pale grey vertical line segments from each point to the grand mean line, representing overall deviation from the grand mean. To the left of the scatterplot is a dotplot of the responses $y_i$, with a triangle indicating the mean. On the far left is a variability meter, initially grey, whose height is proportional to the variance Var($y_i$). Each of these components updates automatically when any of the data points are changed by dragging.

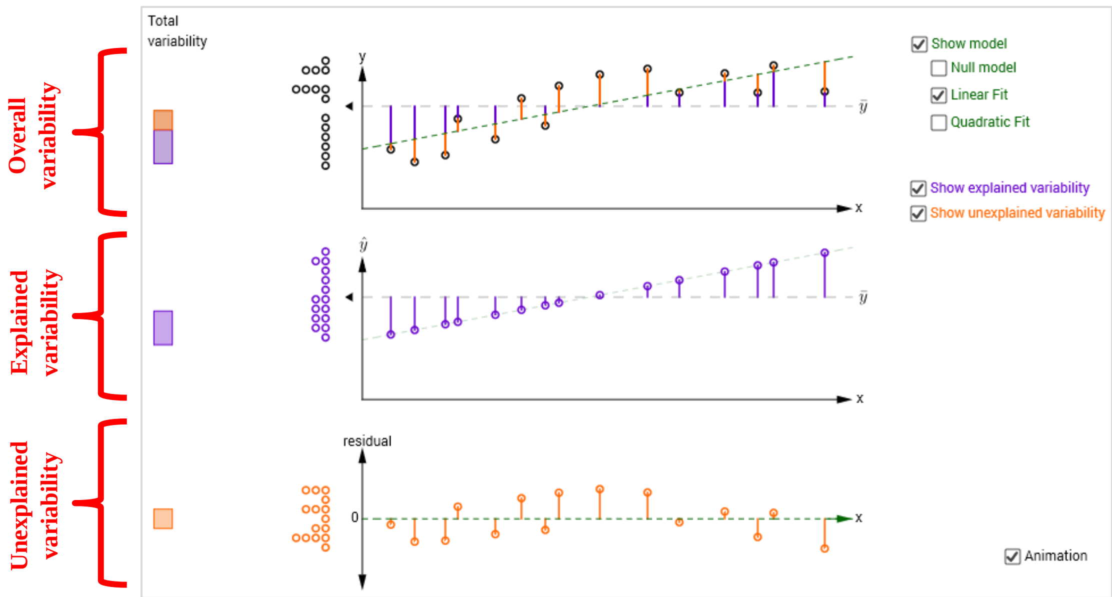



**Figure 2: Screenshot of the regression applet with all components shown and all animations complete.**

Checkboxes to the right allow more elements to be added to the display. The first checkbox, labelled 'Show model', adds a curve representing a regression model to the scatterplot. The user may choose between three options for the regression model: a linear model (of the form $y_i = \alpha + \beta x_i + e_i$, where $\alpha$, $\beta$ are constants and $e_i$ a random residual), a quadratic model (of the form $y_i = \alpha + \beta x_i + \gamma x_i^2 + e_i$), and the null model ($y_i = \bar{y} + e_i$). When the "Show explained variability" checkbox is enabled, the scatterplot points animate and ascend/descend to the grand mean line. At the same time, vertical line segments, coloured purple, appear between the grand mean line and the regression curve. These depict the individual deviations of the predicted values $\hat{y}_i$ about the grand mean $\bar{y}$. Next, copies of the points, grand mean line, and purple segments descend from the original plot to a new set of axes below the original scatterplot, to form a scatterplot of the predicted values $\hat{y}_i$. A dotplot of the $\hat{y}_i$'s appears to the left of the axes, and a variability meter, purple, appears to the far left, representing the variability in the $\hat{y}_i$.

When the final checkbox "Show unexplained variability" is enabled, orange vertical line segments appear between each point and the regression curve, depicting the residuals. The points, grand mean, regression curve, and orange line segments animate and descend from the top plot to a new set of axes at the bottom of the display, to form a scatterplot of the residuals. A dotplot of the residuals appears to the left of the axes, an orange variability meter appears at far left, whose height is proportional to the variance $\text{Var}(e_i)$, and the remaining section of the overall variability meter is coloured. To complete the animation, all visual components reappear in the original scatterplot. Figure 2 shows the applet with all checkboxes enabled and after the animations are finished.

DESIGN CONSIDERATIONS

In both applets, the dotplots show the variability in overall data (top section), variability in means (middle section), and variability about means (bottom section). The variability meters on the left clearly show the partitioning of a total variability into parts corresponding to the explained and unexplained portions. This is a visual representation of the 'partitioning of sum of squares' theorem $SS_{total} = SS_{error} + SS_{treatments}$, and a subtle reinforcement of the independence of the random and deterministic components of the models. The animations show the act of partitioning; they can be repeated over and over to *cement* the abiding image. The movement is 'eye-catching'. However, as in (Wild et al., 2013), we see the animation as valuable when the focus is on understanding the *nature* of partitioning variability, but it may become an undesirable distraction once the nature is understood and the focus shifts to the *effects*. To mitigate this, an additional 'Animation' checkbox allows the lower panels to be shown or hidden without the delay of the animation.

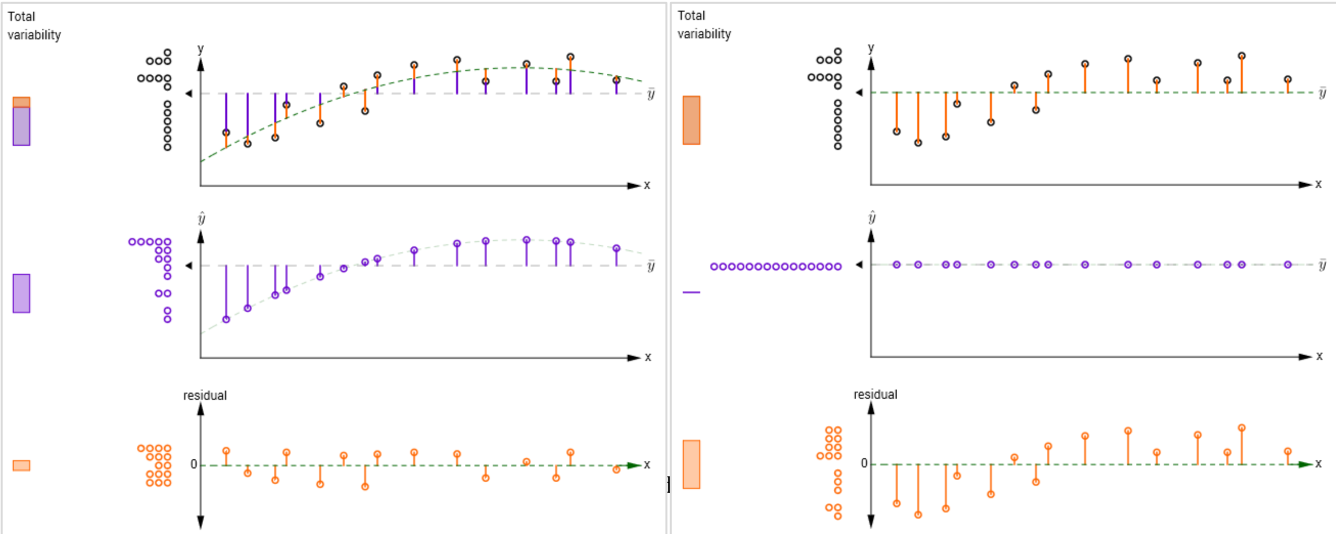
In order to focus on the visuals, we have chosen not to include any numbers or formulas and to use minimal labels in the applets' displays. By omitting a numerical scale or axis label for the variability meters, it is left agnostic whether the variability is measured as a sum of squares, a variance, or perhaps some other measure of variability. This helps to target a more fundamental concept – "the variability can be partitioned" – rather than a specific mathematical realisation of this idea – "the sum of squares can be partitioned". Our aim was a visual representation that would produce *abiding images* of the concept of partitioning variability. Abiding images are encouraged by visuals that are dynamic and *stripped back* so that features of the concept are clearly visible (Wild et al., 2013). The stripping back is about reducing cognitive load (Mayer & Moreno, 2003) and involves the removal of all unnecessary distractions, such as formulae and numbers (in our case), as well as non-essential complexity, such as additional user interface elements to change the sample size or number of groups.

The overall layout of the regression applet mirrors that of the ANOVA applet discussed earlier. The use of a common layout and visual components is intended to help convey the unity of underlying concepts: that regression and one-way ANOVA are two variations of the common idea of modelling data with deterministic and random components. (Indeed, they are two instances of the more general concept of general linear model.) The intent is to "convey the 'sameness' of what is happening [...] to reveal the unity of the underlying way of thinking" (Wild et al., 2013, p. 10).

These are the three features that, we feel, most distinguish these applets from previous applets such as those of CAST and WISE cited earlier: the stripped back visuals for supporting visual reasoning, the use of animations to further make explicit the process of partitioning, and the common layout and visual elements to reinforce the unity of the underlying concepts.

One notable difference between Applets 1 and 2 is the enhanced ability to compare models in the regression applet. By switching between 'linear fit', 'quadratic fit' and 'null model', the user can observe changes in the *nature* of explained and unexplained variability (as seen in the purple and orange line segments comprising the deviation from the mean in the top section), in the *quantity* of 'explained' and 'unexplained' variability (as reflected in the size of the 'explained' and 'unexplained' variability meters or the spread of the dotplots), and in the partitioning of overall variability (as reflected in the break-up of the overall variability meter into purple 'explained' and orange ‘unexplained’ components). This is particularly powerful for the null model, where it is clear from the variability meters and explained/unexplained dotplots that, under this model, ‘all variability is unexplained’. The user can also observe the preservation of total overall variability, as seen by the total size of the overall variability meter remaining unchanged. This can be seen in Figure 2 (a linear fit), Figure 3a (a quadratic fit) and Figure 3b (null model). The data is the same in these three figures; only the choice of model is changed. We note that the ‘null model’ is visible in the ANOVA applet also, in the form of the $\bar{\bar{y}}$ line, though it is not explicitly labelled as such.

Another difference between the two applets is the way that users can manipulate the data. In Applet 1, the user can manipulate the group means and within-group variability, but not the




individual data points. This allows direct manipulation of the explained and unexplained components of variation, and avoids any issues with potential violation of the equal-variances assumption. In Applet 2, the user can drag individual data points directly. This allows the user the flexibility to choose data with a linear, quadratic or other relationship (although it admits the possibility of heteroskedasticity – a trade-off we feel is acceptable).

These applets have been used in an introductory statistics course for first-year science students taught by the second author for several semesters. The majority of students in this course are not specialising in statistics and many do not have advanced mathematical backgrounds. We feel that the emphasis on visuals taken by the applets is appropriate for such students. For some results from an evaluation of the applets (as part of a suite of applets) in this introductory statistics course, we refer the reader to (Morphett et al., 2016).

## MAPPING THE APPLETS INTO A POTENTIAL LEARNING TRAJECTORY

We now illustrate how the applets map to a possible learning trajectory focused on developing an understanding of modelling, which is largely followed in an introductory statistics subject taught by the second author. The learning trajectory for this subject begins by (1) exploring random variation, including sampling variation and other sources of variation. This is followed by (2) modelling variation in outcomes of a random experiment using probability models, eg binomial and normal distributions. (3) With these tools we look at random variation about a single mean. Random variation about the mean is modeled using a distribution. Discussion then moves to competing models, i.e., $\mu = 50$ vs. $\mu \neq 50$, and (informally) the idea of variability that can be explained by the null model versus variability that can’t be explained by the null model. (4) We next move to variability about two means (separate means model). This involves modelling with a deterministic component which depends on the explanatory variable. (5) The next step is variability about several means. We now start talking more formally about variability which is *explained* and *unexplained* - by (the deterministic part of) the model, in each case. (6) We now progress to variability about a more complex model – a regression model – in which the deterministic component of the model is a linear or other function of the explanatory variable. In all steps of this trajectory we are building the notion of variability that can be explained by a proposed model versus variability that is not explained by the model. By steps (5) and onwards, this idea is becoming more formal with the introduction of the analysis of variance ANOVA table.

We first introduce the ANOVA applet at step (5) of this trajectory, however aspects of the applet reflect elements from earlier stages. For instance, sampling variability (1) is seen when generating a new sample, and the 'null model' – shown as the $\bar{\bar{y}}$ line – reflects variation about a single mean (3). We then lead into the regression applet at stage (6), but it also reflects earlier stages, particularly (1) and (3), and is a generalisation of (5). From here, we could lead into hypothesis testing as the comparison of two (nested) models – a possible stage (7).

In this trajectory, the commonality of stages (5)-(7) is seeing ANOVA not just as one single method for comparing several means, but as a powerful and general analysis technique – *analysis* of variance – that involves partitioning variability into 'explained' and 'unexplained' components. The unity of design of the applets, noted earlier, complements this.

We have discussed a pair of applets intended to support visual reasoning about variability and statistical modelling. The design of the applets provides several different ways of viewing the variability in the data. Common visual elements and layout between the applets help re-inforce the unity of underlying concepts, particularly when nested within a curriculum that focuses on statistical modelling, comparison of models and analysis by partitioning variability into explained and unexplained components. In future work, we hope to investigate student use of the applets with regard to the framework of (Reid & Reading, 2010).

On reflection, many of the principles which guided our design of the applets match those discussed in (Wild et al., 2013). The fact that we, largely independently, reached broadly similar considerations as Wild et al when working with different concepts suggests that these principles are sound, robust principles for educational software design in statistics.

ACKNOWLEDGEMENTS

The basic partitioning of variability diagram, and in particular the idea of repeating all sample points at the group mean in the section which illustrates between groups variability, is based on unpublished work by Dr Robert Maillardet, who we would also like to thank for valuable contributions and suggestions during the detailed applet design process.

REFERENCES

Arnold, P., Pfannkuch, M., Wild, C., Regan, M., & Budgett, S. (2011). Enhancing students' inferential reasoning: From hands-on to "movies". *Journal of Statistics Education*, *19*(2).

Dreyfus, T. (1991). On the status of visual reasoning in mathematics and mathematics education. In Furinghetti, F. (Ed.), *Proceedings of the 15th PME International Conference*, Assisi, Italy.

Mayer, R., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, *38*(1), 43-52. doi: 10.1207/S15326985EP3801_6

Morphett, A., Gunn, S., & Maillardet, R. (2015). Developing interactive applets with GeoGebra: processes, technologies. In Blignaut, R. & Kizito, R (Eds.), *Proceedings of Elephant Delta: 10th Southern Hemisphere Conference on the Teaching and Learning of Undergraduate Mathematics and Statistics*, Port Elizabeth, South Africa.

Morphett, A., Maillardet, R., & Gunn, S. (2016). Conceptual Learning with Interactive Applets project report. The University of Melbourne. Available at http://www.melbapplets.ms.unimelb.edu.au/?page_id=201

Reid, J., & Reading, C. (2010). Developing a framework for reasoning about explained and unexplained variation. In Reading, C. (Ed.), *Data and context in statistics education: Towards an evidence based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS-8)*, Ljubljana, Slovenia.

Reid, J., Reading, C., & Ellem, B. (2008). Developing assessment items to measure tertiary students' reasoning about explained and unexplained variability. In Hays, T. & Hussain, R. (Eds.), *Proceedings of the 2nd Annual Postgraduate Research Conference: Bridging the Gap between Ideas and Doing Research*, University of New England, Armidale, Australia.

Wild, C. J., Pfannkuch, M., Regan, M., & Parsonage, R. (2013). Next Steps in Accessible Conceptions of Statistical Inference: Pulling ourselves up by the bootstraps. Unpublished draft, available at http://www.stat.auckland.ac.nz/~wild/TEMP/bootstrap.pdf