Appendix 1. Overview of Bayesian data modelling.

In this brief appendix we describe the methods and present additional detailed results of the Markov chain Monte Carlo modelling used for estimating differences in mean efficacy score for survey responses in which respondents identified their micro-narrative as being associated with particular stages of change (SOC).

In the survey respondents were asked to use a polarity scale that ranged from 0 to 100 and measured efficacy to identify how able to make a difference individuals or groups felt in the response they related. The specific question was:

The individuals or groups in my response were:
Unable to make a difference……Overwhelmed by the opportunities to have an impact.

The resulting efficacy scores were not normally distributed (Figure A1.1), they were also bound between 0 and 100 and had an implicit ideal state in the middle of the score range (i.e. 50). Standard Gaussian assumptions would therefore not be appropriate for these data.

We elected to use a Beta distribution to describe these data given the beta distribution is bound between 0 and 1 and because of its great flexibility in terms of the range of distributional shapes it can represent with different values of its two parameters (α, β) (other polarity distributions in the data set exhibited even more extreme distributional forms than those shown in Figure A1.1).

Preliminary investigation of the data indicated that by using a log transformation and then dividing the result by 5 (to render the results between 0 and 1) the data were more reliably modelled by the Beta distribution.

Using the approach of Gelman et al (1995) we developed a Bayesian hierarchical model to estimate the differences in mean efficacy values across the four SOC. Using R's (R Core Team 2013) capacity to link to JAGS (Just Another Gibbs Sampler)(Plummer 2003, Su and Yajima 2014) we developed a model to estimate the alpha and beta parameters of the (beta) distribution of efficacy scores for each SOC data set (see note 1 below). The model was set up as a hierarchical Bayesian model with three stochastic nodes for estimating mean efficacy for each SOC: the alpha and beta parameters of the *Beta* distribution (with both of these parameters being normally distributed) and the efficacy distribution itself which was a beta distribution. Following Gelman et al (1995, p477) the mean of the efficacy distribution was calculated as: *alpha/(alpha + beta)* and the mode was estimated as *(alpha - 1) / (alpha + beta – 2)*. Once the *alpha* and *beta* parameters were estimated they were used to estimate the mean and mode of the distribution. Differences in means were estimated by subtracting, for each iteration, the value of *mu* (i.e. the mean) of each SOC from the *mu's* of the remaining SOC. Estimates presented in Figure 1a and 1b of the main paper reflect the samples taken from the posterior distributions.

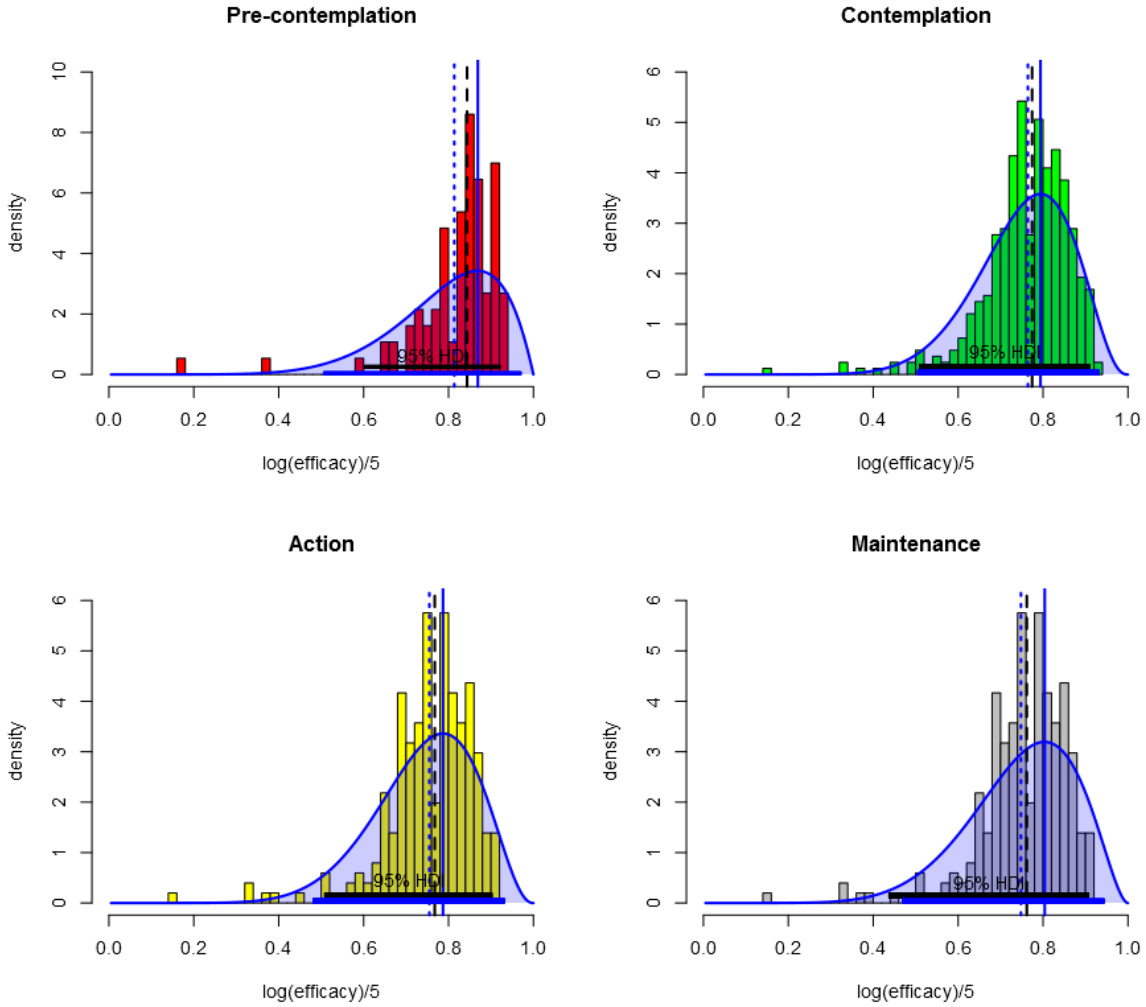

Figure A1.1. Probability density plots of the transformed (y=log(efficacy)/5) efficacy data for each SOC. The shaded curves are the posteriors of the estimated Beta distributions for each SOC. The histograms show the distribution of the transformed data. The solid vertical line is the mode for the Beta distribution. The black dashed vertical line is the median of the data. The dotted blue vertical line is the median of the Beta distribution. The 95% HDI bars at the bottom of each plot identify 95% of the probability mass: the blue bar for the Beta distribution and the black bar for the data associated with each plot.

In each simulation 3 chains were initialised and run for 50,000 iterations. The first 10,000 iterations were discarded. From the remaining 40,000 every tenth result was saved and used for subsequent analyses. For each beta distribution the priors on alpha and beta were initialised at values of 3 and 2 respectively to achieve a prior distribution with a slight negative skew.

The resulting three chains were examined for convergence using density plots of the estimates of each parameter, including the means, and plots of running means for each parameter for each chain. Convergence was excellent for all parameters. The results were also examined for autocorrelation which was negligible over all lags.

Summary results for the model are presented in Table A1.1. The model captured the data distributions reasonably well (Figure A1.1) although in most cases the data distributions were more peaked than the beta although this was only really marked in the pre-contemplation stage (Figure A1.1).

Table A1.1. Summary of the MCMC simulations of the parameters of the Beta distributions (alpha_hyper and beta_hyper) and the mean (mu) and mode of each distribution across SOC. For each parameter, n.eff is a crude measure of effective sample size, and Rhat is the potential scale reduction factor (at convergence, Rhat=1). Suffix ".i" was used for parameters associated with the pre-contemplation stage; suffix ".r" was used for the contemplation / preparation stage; suffix ".c" for the action stage; and suffix ".m" for the maintenance stage.

Parameter group	Parameter	mean	stdev	2.5%	25%	50%	75%	98%	Rhat	n.eff
Parameters of Beta Distributions	alpha.hyper.c	9.464	0.634	8.232	9.046	9.455	9.892	10.728	1.001	11000
	alpha.hyper.i	7.754	0.752	6.299	7.249	7.741	8.252	9.257	1.001	12000
	alpha.hyper.m	8.278	0.652	7.034	7.828	8.266	8.711	9.584	1.001	4300
	alpha.hyper.r	10.575	0.588	9.448	10.169	10.564	10.973	11.739	1.001	12000
	beta.hyper.c	3.292	0.223	2.861	3.141	3.288	3.441	3.735	1.001	12000
	beta.hyper.i	2.023	0.205	1.634	1.884	2.018	2.156	2.44	1.001	12000
	beta.hyper.m	2.782	0.221	2.355	2.633	2.777	2.93	3.228	1.002	3000
	beta.hyper.r	3.488	0.192	3.121	3.358	3.489	3.614	3.879	1.001	12000
Modes of distributions	mode.c	0.787	0.009	0.77	0.781	0.787	0.793	0.805	1.001	12000
	mode.i	0.869	0.016	0.838	0.858	0.869	0.88	0.903	1.001	12000
	mode.m	0.804	0.012	0.78	0.796	0.803	0.811	0.827	1.001	5700
	mode.r	0.794	0.006	0.781	0.79	0.794	0.798	0.806	1.001	12000
Means of distributions	mu.c	0.742	0.007	0.727	0.737	0.742	0.747	0.756	1.001	12000
	mu.i	0.793	0.012	0.769	0.785	0.793	0.801	0.816	1.001	12000
	mu.m	0.748	0.009	0.73	0.742	0.749	0.755	0.766	1.001	12000
	mu.r	0.752	0.005	0.741	0.748	0.752	0.756	0.762	1.001	12000
Model fit	deviance	-1670.1	9.544	-1687.1	-1676.8	-1670.6	-1663.9	-1650	1.001	9600

Notes

1. As the instrument permitted respondents to select more than one SOC there is some overlap in the data sets. Those that told of experiences associated with the pre-contemplation phase did not overlap with any of the other stages.

References cited

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. *Bayesian data analysis*. First edition. Chapman and Hall, London, UK.

Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Pages 20-22 *in* 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Su, Y.-S., and M. Yajima. 2014. R2jags: A Package for Running jags from R.