Social Indicators Research (2006) 79: 369–403
DOI 10.1007/s11205-005-5381-2

BRUCE HEADEY

SUBJECTIVE WELL-BEING: REVISIONS TO DYNAMIC EQUILIBRIUM THEORY USING NATIONAL PANEL DATA AND PANEL REGRESSION METHODS

(Accepted 17 November 2005)

ABSTRACT. This paper partly revises the dynamic equilibrium (DE) theory of subjective well-being (SWB), sometimes termed set point theory. Results from four national panel surveys show that correlations among measures of SWB diminish over time, and that the SWB set points of a minority of individuals substantially change. These results mean that DE theory requires revision to make it more dynamic and enable it to better account for medium term change in SWB. The paper identifies personality traits and life events associated with subsequent changes in SWB. Data come from German, British and Australian panel surveys in which SWB has been measured for between 9 and 20 years. Panel regression random and fixed effects models were used to analyse the data.

KEY WORDS: dynamic equilibrium, life events, personality, SWB

1. INTRODUCTION

A dynamic equilibrium (DE) theory of subjective well-being (SWB) was initially proposed 15 years ago to account for linkages between personality, life events, well-being and ill-being (Headey and Wearing, 1989). Prior to that, Brickman and Campbell (1971; see also Brickman et al., 1978) had shown that people usually return to a baseline – or equilibrium level, or set point – of happiness following even such major life events as becoming a paraplegic and winning a large sum in a lottery. DE theory is now sometimes labelled set point theory. This is perhaps a clearer label, or a more immediately understandable description of one key aspect of the theory, but in itself the observation that individuals generally seem to oscillate around their own set point of SWB is just that; an observation not an explanatory theory.

DE theory has been extended in two main ways in recent years. Evidence has been adduced about additional stabilizing factors which tend to keep people close to their happiness/SWB set point. Headey and Wearing (1989) attributed long-term stability to the stable traits of extraversion (E) and

neuroticism (N). Lykken and Tellegen (1996), using the Minnesota Twin Study, showed that heredity generally (and not just E and N which are substantially heritable), is a powerful influence on lifetime SWB (see also Lykken, 1999). Other researchers have investigated the effects of life events which can cause medium term and perhaps permanent change in set points. These events include the unexpected death of a child (Wortman and Silver, 1987), repeated spells of unemployment which have a 'scarring effect' (Clark et al., 2004), becoming widowed and perhaps getting married (Lucas et al., 2003). The only favorable or positive 'event' unambiguously shown to enhance long-term SWB seems to be cosmetic surgery (Wengle, 1986; Frederick and Loewenstein, 1999).

DE theory remains controversial partly because the two lines of research just described pull in different directions, although they are not necessarily contradictory. The first elucidates set-point stabilizers, while the second directs attention to destabilizers. Clearly, if it were common for destabilizers to overwhelm stabilizers, then dynamic equilibrium/set point theory would not hold. The theory depends on finding that most people, most of the time, have stable levels of SWB. Some recent papers, using longitudinal data, report medium term changes in SWB but stop short of concluding that DE/set point theory requires revision (Lucas et al., 2003; Clark et al., 2004; Fujita and Diener, 2005).

The purpose of this paper is to revise DE theory, using data from four national socio-economic panel (i.e. longitudinal) surveys: the West and East German (SOEP) panels, the British (BHPS) panel and the Australian (VQOL) panel. The main outcomes are: (1) clear evidence that DE theory requires some revision; evidence centered on finding that panel studies show that the stability of life satisfaction diminishes slowly over time (hypothesis 1 below), (2) confirmation of a key aspect of the theory relating to the claim that life events are partly endogenous (driven by person characteristics) and, to that extent, do not change SWB (hypotheses 5.1 and 5.2 below), and (3) preliminary analysis of the factors associated with medium term upward and downward changes in SWB in relatively small but not tiny minorities of the population.

These new results show that previous claims that the stability of SWB scarcely diminishes over time were incorrect (Headey and Wearing, 1989, 1992). Even less correct was the claim that SWB is so strongly hereditary that attempts to enhance it are doomed to fail (Lykken and Tellegen, 1996). The effect of the revisions proposed here is to make dynamic equilibrium theory more dynamic. We point to 'starting characteristics' (including personality traits) which predispose people to experience more rather than less

change in SWB. Also considered are life events associated with change in SWB.

In analyzing data from the four panels, some use will be made of *panel regression analysis*; a group of Generalized Least Squares techniques developed by econometricians to explore the dynamics of change in both macro and micro panel data (Lillard and Willis, 1978; Baltagi, 1995). Panel regression analysis has been little used in SWB research, partly because of lack of panel data and partly because the techniques have only recently become available in easy-to-use software (but see Clark et al., 2004; Frijters et al., 2004).

A final introductory point: in this paper DE theory is taken to apply not just to *well-being* (life satisfaction and positive affects), but also to *ill-being* (anxiety states and related negative affects). The theory as initially put forward related to both well-being and ill-being (Headey and Wearing, 1989, 1992), and it is suggested that this was and is a valuable aspect. In recent years the tendency has been to refer solely to a well-being or happiness set point. It will be assumed that well-being (henceforward WB) and ill-being (IB) have been shown to be distinct dimensions, which are only moderately negatively correlated, and not just opposite ends of the same dimension (Bradburn, 1969). This is not accepted by all researchers in the field, but by the large majority (Headey and Wearing, 1992; Argyle, 1987, 1999; Diener et al., 1999).[1] Physiological evidence backs the majority view, indicating that the centers of the brain registering pleasure/positive affect and pain/negative affect are separate, although they interact as 'opposing systems' (Kahneman, 1999).

1.1. *DE Theory/Set Point Theory Reformulated*

A helpful way to understand DE theory is to compare it with human capital theory in economics. Human capital theory (Becker, 1975) proposes that each individual has a predictable lifetime income curve; predictable on the basis of his/her human capital (education, relevant work experience, and intangible capital like entrepreneurial skills). Most people's incomes rise during their working lifetime up to the age of about 50, and then decline in later working years and retirement. It is clear that annual incomes fluctuate due to random shocks, but the prediction is that most people revert to long-term trend. A few rise from rags to riches and a few fall from riches to rags, but the existence of these small minorities is not seen as total refutation of human capital theory.

DE theory is similar but simpler. It predicts not a lifetime curve but fluctuations around a straight line in SWB; each individual has WB and IB set points. An individual's WB and IB fluctuate from year to year, due to the fact

that events and experiences are partly exogenous (as well as partly endogenous), but will then return to baseline. The analogues of human capital in SWB theory are the personality traits E, N and O. By analogy, too, it should not be seen as complete refutation of DE theory if small minorities register large gains or losses in SWB.

DE theory was initially put forward to account for two observations in the Australian (VQOL) Panel Survey. The first had been made before (Brickman and Campbell, 1971), the second seemed somewhat new and unexpected.

1. *Some people were persistently happier than others*. To be more precise, some respondents rated consistently higher on measures of WB than others, and some rated consistently higher on measures of IB. Measures of WB and IB were only moderately negatively correlated, so as well as observing individuals who were high on WB and low on IB, and vice-versa, we also found people who were high on both WB and IB, and a fourth group who were low on both WB and IB.
2. *The same life events and experiences kept happening to the same people*. This was a key surprise result in the Australian panel. It provided the clue that personality traits, life events, WB and IB might be linked; they might be 'in dynamic equilibrium'.

1.2. *Hypotheses*

DE theory accounts for these observations, using the following building blocks: personality (N, E and O – Openness to Experience), life events (positive and negative), anticipatory and adaptive mechanisms invoked by life events, and WB and IB. These building blocks are linked in six sets of hypotheses, which are laid out in more detail than in earlier papers (Headey and Wearing, 1989, 1992) and reassessed using new evidence and methods. A further new set of hypotheses, aiming to account for medium term changes in WB and IB, are better left to the beginning of Section B, which is entirely about change.

Hypothesis 1: Each person has his/her own stable equilibrium levels (or set points) of WB and IB.

This hypothesis generates the following specific predictions: (i) representative samples of individuals followed over time will have stable aggregate levels of WB and IB (ii) *within-person* over time coefficients of variation in WB and IB will be significantly smaller than *across person* coefficients of variation, and (iii) over time most individuals will change little from their

baseline (i.e. starting) levels of WB and IB (iv) the over-time correlations of WB and of IB will be high, and will *not* diminish as the length of the period of measurement increases.

Prediction (i) follows clearly from the hypothesis; if all individuals have stable levels of WB and IB, then a representative panel, followed over time, will have stable means and standard deviations. Prediction (ii) says that individuals in the national panels, who start out at different levels of WB and IB, will then continue at much the same levels over time, exhibiting significantly less variation than is found if we compare across individuals at any given moment in time (see also Fujita and Diener, 2005). Prediction (iii) is really just a restatement of the hypothesis, setting up baseline measures of WB and IB and then predicting that individuals will not subsequently deviate much from their own baselines. Prediction (iv) is a very strict test of Hypothesis 1. It says that if WB and IB really are stable, then the correlation between, say, time 1 and time 5 measures should be no lower than between time 1 and time 2 measures (despite the fact that correlations among almost all other psychological measures diminish over time).

Hypothesis 2: Levels of WB and IB depend partly on E and N. People who rate high on E and low on N have high levels of WB and low levels of IB (Quadrant 1). People who score low on E and high on N rate low on WB and high on IB (Quadrant 2). People who score high on both E and N also score high on both WB and IB (Quadrant 3). People who rate low on both E and N score low on both WB and IB (Quadrant 4).[2]

Hypothesis 3.1: Each person has a tendency to display repeating patterns of life events and experiences. Positive (favourable,) events 'scores' (i.e. frequency of experiencing positive events) are correlated over time, as are negative (adverse) events 'scores'.

Hypothesis 3.2: Positive and negative events are also correlated over time. That is, the more positive events a person experiences, the more negative events he/she is also likely to experience (see Hypothesis 4.3 below for explanation of this apparently counter-intuitive hypothesis).

Hypothesis 4: The repeating patterns of events which people experience are driven by three personality traits – N, E, O – and stage of the life cycle.

H4.1: People who experience many positive events and few negative events score high on E, low on N, high on O and are relatively young.

H4.2: People who experience few positive events and many negative events score low on E, high on N, high on O and are relatively young.

H4.3: People who experience many positive and many negative events score high on E, high on N, and high on O and are relatively young.

H4.4: People who experience few positive and few negative events score low on E, low on N, low on O, and are relatively old.

This fourth set of hypotheses, if confirmed, would suggest that people have their own individual 'normal' patterns of life events. Reasons for the proposed links between E and N, on the one hand, and positive and negative events on the other, are fairly obvious and similar to the reasons linking these traits to WB and IB (Costa and McCrae, 1980a). The importance of O in this context is that it appears that 'openness to experience' is linked to experiencing more events of all kinds and hence to the observed positive correlation (across persons) between positive events scores and negative events scores. It should also be noted, as a relatively uninteresting part of hypotheses 4.1–4.4, that age is also implicated. Younger people are at a stage of the life cycle when more major events (marriage, divorce, promotion, unemployment, etc) are likely to happen than is the case with older people. It is also the case that women report slightly more positive and negative events than men (Magnus et al., 1993), and slightly higher positive affect and negative affect (Andrews and Withey, 1976).

Hypothesis 5.1: To the extent that, in any given time period, a person just repeats the pattern of positive and negative life events that is normal (mean level) for him/her, then WB and IB will remain at or revert to their set points.

Hypothesis 5.2: To the extent that, in a given time period, a person deviates from his/her own normal pattern of events (i.e. relatively exogenous events happen), then WB and/or IB will deviate from their set points.

Hypothesis 6: Major life events are usually anticipated beforehand and habituated to afterwards. These adaptive mechanisms reduce their impact in changing levels of WB and IB and promote return to set points.

2. METHODS

Data come from the only four panel studies which, so far as we know, have measured either WB or IB (or both) annually for a decade or longer. In this paper the West and East German (SOEP) panels are used when we test hypotheses which just require long-running data on WB. The British (BHPS) panel is used when we require long-running data on IB. The Australian (VQOL) panel, conducted in the State of Victoria, was the only one designed specifically to investigate SWB. It is used when hypothesis-testing requires data on personality and life events, as well as WB and IB.

2.1. *Samples: SOEP (West Germany), SOEP (East Germany), BHPS (Britain) and VQOL (Australia)*

The West German segment of the SOEP panel is the longest-running panel in the world to collect data on WB (life satisfaction). It began in 1984 with a sample of 12,541 respondents. Interviews have been conducted annually ever since. Everyone in the household aged 16 and over is interviewed. The representativeness of the panel is maintained by interviewing 'split-offs' and their new families. So when a young person leaves home ('splits off') to marry and set up a new family, the entire new family becomes part of the panel. The main sample has also been boosted by the addition of new immigrant samples, a special sample of the rich, and recruitment of new respondents partly to increase numbers in 'policy groups'. The main topics covered are family, income and labor force dynamics, but a question on life satisfaction has been included every year. The data used in this paper mainly relate to the 3340 respondents who reported their level of life satisfaction every year for 18 consecutive years from 1985 to 2002.[3]

An East German panel was added to SOEP in 1990 after the Berlin Wall came down but before German reunification; the main aim then was to get a picture of social and economic conditions in the GDR before it disappeared. From 1991 onwards the East sample became part of the 'normal' SOEP, so in this paper we use data for 1991–2002. The sample size in the East was initially 4453 and it has not been boosted since, except by split-offs. A total of 1012 individuals reported their life satisfaction every year. The East sample is particularly valuable for investigating the stability of SWB. During the 1990s panel members were living in a 'country' going through massive social, political and economic changes as it moved towards democracy and capitalism. Income inequality increased a great deal, income mobility was exceptionally high, unemployment rose rapidly, and so did the divorce rate (Krause, 2003). The pertinent research question here is, 'Was life satisfaction stable even in these circumstances?' (see also Frijters et al., 2004). If it was, then we have compelling evidence that it will nearly always be stable.

The British (BHPS) panel also began in 1991 with a sample of 10,300 respondents. As in the SOEP, all household members aged 16 and over are interviewed, and interviews are annual. Representativeness is maintained by following split-offs. Again, the main focus is economic. For present purposes, the special value of the BHPS is that it includes an annual measure of IB, namely the General Health Questionnaire (GHQ; see below).[4] The data

used here are for the 4223 respondents who completed the GHQ in 1991–2002.

The Australian (VQOL) panel ran from 1981 to 1989, with interviews being conducted every 2 years (Headey and Wearing, 1992). The initial sample size was 942 with one person per household being interviewed. By 1989 the usable sample was 477. The Australian panel is much smaller than the others but was designed specifically for research on SWB. So, as noted above, it includes data on personality and life events, as well as WB (life satisfaction) and IB (negative affect). It is not suggested that the Australian panel, unlike the others included here, is a representative sample; but this is not important, given the focus on what are claimed to be universal linkages between personality, events, well-being and ill-being.

Except where specifically mentioned, all analyses use *balanced panels;* that is, only respondents who answered SWB questions in every wave of the panel are included. The reason for using a balanced rather than an unbalanced panel design is that tests of hypotheses about the stability of SWB require continuous data. Also, we do not use sample weights (neither cross-sectional nor longitudinal weights) because the aim is to obtain results relating to stability and change in the individuals sampled, not make population estimates.

2.2. *Measures*

Two types of measures of SWB are collected by researchers: global self-reports and 'on-line' (experience sampling) measures. All the measures here are global self-reports; survey items requiring respondents to make summary judgments about their 'life satisfaction', moods or other mental states. It is acknowledged that advances in SWB theory may well require us also to collect 'on-line' measures (Kahneman, 1999), but for now only global self-reports are available on a multi-year basis.

The dependent variables are measures of WB and IB. WB is measured by a 0–10 scale of life satisfaction in the German panels, and a 1–9 Delighted-Terrible scale (Andrews and Withey, 1976) in the Australian panel. The life satisfaction scale is a single item on which 0 means 'very dissatisfied' and 10 means 'very satisfied'. Clearly, single item scales are not the most valid measures of WB available, but they have been reviewed as having reasonably adequate reliability and validity (Diener et al., 1999). The Andrews and Withey Delighted-Terrible scale involves asking respondents, 'How do you feel about your life-as-a-whole?', using a response scale running from 1 ('terrible') to 9 ('delighted'). The item was asked twice about 20 min apart in

each wave of the Australian survey. Answers were averaged to give a life satisfaction index. The index had a Cronbach α reliability of 0.83 in 1987 and satisfactory convergent validity with Diener et al.'s (1985). Satisfaction With Life Scale and other widely used measures of WB (Larsen et al., 1985; Headey et al., 1993). There appears to be no problem in using the life satisfaction and Delighted-Terrible measures as more or less comparable measures of well-being. Both combine cognitive and affective components of well-being, with the life satisfaction scale perhaps being the more cognitive measure (Andrews and Withey, 1976; Diener et al., 1985).

The measure of IB in the British panel was the 12-item GHQ, which was designed for use by health practitioners as a screening instrument for current mental health problems (Goldberg, 1978). It was intended as a state rather than a trait measure with typical items being 'felt constantly under strain' and 'lost much sleep over worry'. The items were asked in the British panel on a 4-point scale with the end points labeled 'more than usual' and 'much less than usual'. Scores were averaged to construct the GHQ index. The GHQ has a Cronbach α of 0.90 and reasonably satisfactory convergent validity with other IB scales including the Bradburn Negative Affect index and the Beck Depression Inventory (Headey et al., 1993).

In the Australian panel the measure of IB was the Bradburn (1969) Negative Affect scale, a 5-item index comprising 'yes–no' questions on the lines of, 'During the last few weeks did you ever feel very lonely or remote from other people', and 'so restless you couldn't sit long in a chair?'. The index had a Cronbach α of 0.65 and a correlation with the Watson et al. (1988) negative affect scale of 0.54.

The explanatory variables used in the Results section nearly all come from the Australian panel. They relate to personality and life events. The main personality scale was Form B of the Eysenck Personality Inventory (Eysenck and Eysenck, 1964). E and N are both measured by 24 'yes–no' items. In the 1987 survey the Costa and McCrae (1985) 'openness to feelings' scale – one of the facets of O – was added. This had a Cronbach α of 0.76.

The life events inventory in VQOL was modified from the Henderson et al. (1981) inventory which was itself based on Holmes and Rahe's (1967) innovative instrument. It comprised 93 life events (some putatively positive and some negative) which might have happened to respondents since last interview. Henderson et al. (1981) reported a 2-week test–retest reliability of 0.89. The validity of life events inventories has been questioned on the grounds that reports of some events, especially relatively subjective ones (e.g. 'trouble with people at work'), are biased by personality traits, particularly N, by current mood, and by retrospective bias (Diener et al., 1984;

Schroeder and Costa, 1984). However, it appears that reports of relatively objective events (e.g. 'became unemployed') are only slightly affected (Maddi et al., 1987; Rowlison and Felner, 1988).

The usual way to score life events inventories is to calculate a total stress or life change score based on predetermined weights. Given our focus on WB and IB, it was more appropriate to distinguish between putatively positive and negative events. We constructed separate positive and negative events scores with each event having a unit weight. Like previous researchers we found that unit weighting enabled us to account for as much variance in dependent variables as more elaborate stress scores (Rahe and Arthur, 1978).

For ease of interpreting results, and especially the size of regression effects, all variables and indices with arbitrary metrics were transformed to run from 0 to 10. The German life satisfaction measures were on that scale already, but the Australian life satisfaction, the GHQ, Negative Affect and personality trait measures were adjusted. Variables with non-arbitrary metrics (e.g. gender, age and life events) were not adjusted.

2.3. *Panel Regression Analysis*

As in much previous research, starting with Andrews and Withey (1976), it was found that regression analysis yielded essentially the same results as ordinal techniques, even though the dependent variables are strictly speaking ordinal, rather interval or ratio scale measures. However, it is often unsatisfactory to use ordinary least squares (OLS) regression for analyzing panel data, because the assumption that the error terms in successive waves of data are uncorrelated with right-hand side (RHS) variables may not be met. The assumption can be formally tested using the Breusch and Pagan Lagrange multiplier test (1980). So, in analyzing panels, econometricians usually prefer to use Generalized Least Squares (GLS) 'effects' models; either random effects or fixed effects (Baltagi, 1995). In deciding between random and fixed effects, one needs both to take account of theory and also analyze the errors (residuals) in the multi-year equations (Lillard and Willis, 1978). If errors are random over time, then a random effects model is likely to be appropriate. If non-random patterns are detected, then a fixed effects model may be preferred. However, theory may override statistical concerns. If one is testing a 'between-person' hypothesis, it may be preferable to stick to a random effects model (i.e. essentially a between-persons model). Conversely, if one is testing what seems essentially a 'within-person' hypothesis, the case for a fixed effects model may be overwhelming.[5]

In this paper our concern is with errors potentially due to unobserved (unmeasured) individual characteristics; also commonly referred to as unobserved heterogeneity. A typical pattern in panel data on SWB is to find that the some individuals repeatedly have positive residuals – and others have negative residuals – when one tries to predict their SWB. A probable (virtually certain?) explanation is that we have not measured all the hereditary factors and personality traits which could affect outcomes. Standard statistical theory tells us that if any of these omitted variables are correlated with variables on the RHS, then regression estimates are potentially biassed. The presence of statistically significant unobserved fixed effects can be detected by the Hausman test (1978).

Last, it is worth noting that one of the strengths of panel regression analysis is that it can make use of information about change between each pair of time points, and not just between the start and end points of a panel survey. Some previous analyses of SWB panel data have run the risk of giving biased results due to focusing only on start and end points.

3. RESULTS

In Section A the hypotheses set out above are tested. Section B is exploratory and attempts revisions and extensions to DE theory.

3.1. *Section A: Testing Hypotheses Derived from DE Theory*

Hypothesis 1: Each person has his/her own stable equilibrium levels (or set points) of WB and IB.

In previous papers, using the Australian panel data, we tested the four specific predictions derived from this hypothesis discussed above, and all seemed to be confirmed (Headey and Wearing, 1989, 1992). However, as we are about to see, data from the newer longer-running panels confirm the first two predictions, yield ambiguous results for the third, and clearly falsify the fourth. The predictions are tested in Panels A–D of Table I. Note that in Panels A, B and D all measures of WB (life satisfaction) and IB (the GHQ) have been combined into 3-year blocks. Averaging scores into 2 or 3-year blocks (rather than being at the mercy of annual fluctuations) has the benefit of making over-time trends much clearer and is one reason why it is possible to see that the earlier interpretation of the Australian panel results was partly wrong.

TABLE I

Stability and change in WB and IB 1985–2002: West Germany ($N = 3340$), East Germany ($N = 1012$) and Britain ($N = 4223$)

	W. Germany (1985–2002)						E. Germany (1991–2002)				Britain (1991–2002)			
Panel A: Aggregate stability of WB and IB: Means[a]														
	8587	8890	9193	9496	9799	0002	9193	9496	9799	0002	9193	9496	9799	0002
Life satis. (0–10)	7.3	7.2	7.3	7.1	7.0	6.9	6.1	6.3	6.4	6.3	na			
GHQ (0–10)	na						na				3.0	3.1	3.1	3.1
Panel B: Within-person versus across-person coefficients of variation														
	Within-person		Across-person[b]		Significant level?		Within-person	Across-person[b]	Significant level?		Within-person	Across-person[b]	Significant level?	
Life satis. (0–10)	0.19		0.25		0.0001		0.20	0.27	0.0001		na			
GHQ (0–10)	na						na				0.32	0.44	0.0001	
Panel C: Within-person changes from baseline (%)														
Baseline = mean of first 3 years of the panel	+3 years	+6	+9	+12	+15		+3 years	+6	+9		+3 years	+6	+9	

Life satis. (0–10)											
% Who changed < 1 point	68.2	63.4	59.3	56.3	54.2	70.5	62.9	61.0	na		
% Who changed > 1 & le 2	21.8	24.1	24.7	25.0	26.3	22.6	25.1	25.4			
% Who changed > 2 points	9.7	12.5	16.0	18.7	19.5	6.9	12.0	13.6			
	(100)	(100)	(100)	(100)	(100)	(100)	(100)	(100)			
GHQ (0–10)											
% Change < =1 point	na					na			75.7	71.2	70.2
% Change > 1 and < =2 points									18.4	20.5	21.7
% change> 2 points									5.9	8.2	8.1
									(100)	(100)	(100)
Panel D: Pearson correlations of WB and IB over time											
Baseline = mean of first 3 years of the panel	+3 years	+6	+9	+12	+15	+3 years	+6	+9	+3 years	+6	+9
Life satis. (0–10)	0.64	0.53	0.47	0.44	0.41	0.66	0.55	0.51	na		
GHQ (0–10)	na					na			0.62	0.51	0.51

[a]Standard deviations remained very stable from year to year. In West Germany the standard deviation of life satisfaction ranged only from a low of 1.74 (2001) to a high of 2.0 (1985). In East Germany standard deviations ranged from 1.66 (1997) to 1.95 (1991). In Britain the standard deviations of the GHQ ranged from 1.28 (1991) to 1.48 (2002).

[b]Mean of 18 annual coefficients of variation for W. Germany, and mean of 12 for E. Germany and UK.

Panel A of Table I appears to confirm our first hypothesis predicting a high degree of aggregate stability in WB and IB (but plainly this does not demonstrate individual level stability; see below). In West Germany there has been just a small decline in life satisfaction during the long period of economic stagnation from 1993 onwards, when the reunification boom ended. Even this small decline may be over-stated due to a measurement artifact; the tendency of respondents to report higher WB in the first waves of a panel than in later years (Frijters et al., 2004). In East Germany life satisfaction has oscillated only between 6.1 and 6.4 on the 0–10 scale, despite the massive social and economic changes referred to earlier. In Britain, the measure of IB – the GHQ – has remained almost stationary on 3.0 or 3.1.

Turning to the second hypothesis (Panel B), it is plain that in all three countries within-person coefficients of variation in life satisfaction were significantly ($p < 0.0001$) lower than across-person coefficients; over 20% lower in W. Germany and about 26 % lower in E. Germany. In the British panel a similarly large difference was found between the mean within-person coefficient for the GHQ and the mean of annual across-person coefficients. The average within-person coefficient for 1991–2002 was about 27% lower than the mean of annual across-person coefficients.

The results in Panel C give the first clue that the stability of both WB and IB, although remaining high, nevertheless diminish over time. The average score of each survey respondent on life satisfaction (or the GHQ) in the first 3 years of the panel provides a baseline. Panel C then shows by how much respondents changed from baseline after 3, 6, 9 years and so on. It can be seen that the percentage whose scores shifted by 1 point or less became smaller the more years were taken into account, and correspondingly the percentage whose scores shifted by over 2 points became larger. It is also pertinent that in the West German panel, where 18 years of data are available, the percentage who changed by more than 2 points was considerably larger by 2000–2002 than in the other two panels, where only a 12-year period was available. It is worth noting that the decline in stability of life satisfaction appears to be greater than for the GHQ. A similar result was obtained in the Australian panel, in which the measure of IB was the Bradburn Negative Affect scale rather than the GHQ (Headey and Wearing, 1989). Whether this means that WB is less stable than IB must be regarded as uncertain on the basis of this limited evidence, but it is a possibility.

Panel D, which shows diminishing correlations over time, decisively confirms the trend towards diminishing stability. In West Germany the correlation between respondents' life satisfaction scores in the baseline period (1985–1987) and the last 3 years (2000–2002) fell to 0.41. In East

Germany, for the shorter time period, the correlation between baseline and 2000–2002 was 0.51, and the same correlation was found in Britain for the GHQ.

These diminishing correlations are crucial evidence showing that DE theory – or set point theory – requires revision. They indicate a clear need to account for why, despite high overall stability, a significant minority of people record substantial change in WB and/or IB. We return to this issue in Section B of the paper.

Hypothesis 2: Levels of WB and IB depend partly on E and N. People who rate high on E and low on N have high levels of WB and low levels of IB (quadrant 1). People who score low on E and high on N rate low on WB and high on IB (quadrant 2). People who score high on both E and N also score high on both WB and IB (quadrant 3). People who rate low on both E and N score low on both WB and IB (quadrant 4).

Because this hypothesis links personality to both WB and IB, only the Australian panel is available to test it. It has been known for a long time that E and N are correlated with WB (Wilson, 1967), but their differential relationship with IB has been less well understood (but see Costa and McCrae, 1980b). Table II gives Pearson correlations and also reports results on WB and IB for *four quadrants* of respondents split at the mean on their E and N scores. For all variables the mean of respondents' scores during the life of the panel was used, with respondents not interviewed at least three times being omitted.[6]

As is well known, E correlates moderately with WB ($r = 0.28$) but only slightly and negatively with IB (-0.10). N, on the other hand, is strongly related to IB ($r = 0.54$), but less so to WB ($r = -0.29$). Because E and N are themselves correlated only -0.18, it is possible to divide respondents into four roughly equal quadrants in terms of their personality and WB/IB profile.[7]

In Panel B of Table II it can be seen that the four quadrants are quite distinct from each other in terms of both life satisfaction and negative affect. Respondents in Quadrant 1 (high E, low N) have the best of it in terms of both WB and IB. They score significantly higher on life satisfaction than the other three Quadrants, and also rate significantly lower on negative affect than all but Quadrant 4. Respondents in Quadrant 2 are in the worst condition. They rate lower on life satisfaction than people in the other quadrants and report significantly higher levels of negative affect than all but Quadrant 3. Quadrants 3 and 4 are the two intermediate groups in terms of both WB and IB. However Quadrant 4 could be said to be doing better

TABLE II

Personality traits E and N related to WB and IB: Correlations and differences among four quadrants

Panel A: Correlations (N = 602)				
Extraversion	1.00			
Neuroticism	−0.18	1.00		
Life satisfaction	0.28	−0.29	1.00	
Negative affect	−0.10	0.54	−0.49	1.00

Panel B: Four quadrants – differences in life satisfaction and negative affect				
	Life satis. (0–10)		Neg. Affect (0–10)	
	Mean (SD)	Significant level?	Mean (SD)	Significant level?
Quadrant 1 vs. Quadrant 2	7.9 (1.0)	$p < 0.001$	1.6 (1.3)	$p < 0.001$
	7.0 (1.2)		3.3 (2.1)	
Quadrant 1 vs. Quadrant 3	7.9 (1.0)	$p < 0.001$	1.6 (1.3)	$p < 0.001$
	7.5 (1.2)		3.3 (2.1)	
Quadrant 1 vs. Quadrant 4	7.9 (1.0)	$p < 0.0$l	1.6 (1.3)	ns
	7.5 (1.0)		1.8 (1.6)	
Quadrant 2 vs. Quadrant 3	7.0 (1.2)	$p < 0.001$	3.3 (2.1)	ns
	7.5 (1.2)		3.3 (2.1)	
Quadrant 2 vs. Quadrant 4	7.0 (1.2)	$p < 0.001$	3.3 (2.1)	$p < 0.001$
	7.5 (1.0)		1.8 (1.6)	
Quadrant 3 vs. Quadrant 4	7.5 (1.2)	ns	3.3 (2.1)	$p < 0.001$
	7.5 (1.0)		1.8 (1.6)	

Notes: Quadrant 1 = high E, low N; Quadrant 2 = low E, high N; Quadrant 3 = high E, high N; Quadrant 4 = low E, low N. Quadrant Ns: Q1 = 182, Q2 = 157, Q3 = 135, Q4 = 128. ns = not significant at the 0.05 level.

than Quadrant 3 because, although life satisfaction scores are much the same, Quadrant 4 rates lower on negative affect.

Hypothesis 3.1: Each person has a tendency to display repeating patterns of life events and experiences Positive events 'scores' are correlated over time, as are negative events 'scores'.

Hypothesis 3.2: Positive and negative events are also correlated over time. That is, the more positive events a person experiences, the more negative events he/she is also likely to experience.

Table III gives over-time correlations for life events from the Australian panel. The top half of the table gives result for all events in the inventory which could reasonably be classified as 'positive' or 'negative' (listed in Headey and Wearing, 1992). The second half of the table is intended to deal with the reasonable objection that the correlations could be due to selective perception. That is, some people are known to be predisposed to recall positive events, while others are predisposed to recall negative events (Schroeder and Costa, 1984). This may be linked to the tendency of extraverts to 'magnify' and respond more favourably to positive events than introverts, and the tendency for those who score high on N to 'magnify' and respond more negatively than low N scorers to negative events (Diener and Lucas, 1999). So, in the upper right part of Table III (in italics), analysis is restricted to events which appear so 'objective' that recall bias would be minimal.[8]

TABLE III

Over-time correlations among positive events (PE) and negative events (NE)

	PE83	PE85	PE87	PE89	NE83	NE85	NE87	NE89
PE83	1.00	*0.38*	*0.32*	*0.22*	*0.13*	*0.11***	*0.14***	*0.11**
PE85	0.53	1.00	*0.35*	*0.25*	*0.10**	*0.10***	*0.13***	*0.13***
PE87	0.43	0.50	1.00	*0.35*	*0.04*^{ns}	*0.07*^{ns}	*0.07*^{ns}	*0.12***
PE89	0.35	0.44	0.50	1.00	*0.04*^{ns}	*.05*^{ns}	*0.15***	*0.17*
NE83	0.16	0.18	0.20	0.11*	1.00	*0.35*	*0.18*	*0.24*
NE85	0.21	0.28	0.32	0.23	0.49	1.00	*0.33*	*0.26*
NE87	0.24	0.32	0.44	0.28	0.33	0.49	1.00	*0.29*
NE89	0.21	0.27	0.32	0.38	0.38	0.46	0.47	1.00

Key: PE83 = Positive Events occurring in 1981–1983, PE85 = Positive Events in 1983–1985, etc. Similarly, NE83 = Negative Events occurring in 1981–1983, NE85 = Negative Events in 1983–1985, etc.

Note: All correlations were significant at the 0.001 level, except those shown with **Significant at 0.01; *significant at 0.05; ns, not significant at 0.05.

It can be seen that all over-time correlations among positive events are statistically significant at the 0.001 level, as are over-time correlations among negative events. When the list is restricted to 'objective' events, the correlations are smaller but still significant at the 0.001 level. Other investigators, using a longer list of 'objective' events, have replicated these results (Magnus et al., 1993).

Notice, too, that positive events scores are *positively* correlated with negative events scores.[9] This suggests that there must some people to whom lots of events of all kinds happen, and others to whom little happens.

A reasonable and indeed almost unavoidable inference from the evidence in Table III is that the events which happen to people must be partly endogenous; that is, driven by their own more or less stable characteristics.[10] Clearly, if events were entirely exogenous – if they were random shocks – then events scores would be uncorrelated over time.[11] So what causes the correlations?

Hypothesis 4: The repeating patterns of events which people experience are driven by three personality traits – E, N, O – and age.

H4.1: People who experience many positive events and few negative events score high on E, low on N, high on O and are relatively young.

H4.2: People who experience few positive events and many negative events score low on E, high on N, high on O and are relatively young.

H4.3: People who experience many positive and many negative events score high on E, high on N, high on O and are relatively young.

H4.4: People who experience few positive and few negative events score low on E, low on N, low on O and are relatively old.

First, to test the hypothesis that repeating patterns of events are driven by E, N, O and age, Table IV gives Pearson correlations and regression equations. Then, Table V shows events scores for the four Quadrants. Each person's scores in these tables are his/her mean score for a minimum of three waves of the Australian panel.

The correlations here are only moderate, but they clearly show that repeating patterns of events are partly driven by E, N and O. E drives only positive events, while N drives negative events. Interestingly, O correlates 0.25 with both positive and negative events. From the standpoint of personality theory, it is plausible that 'openness to experience' would be positively related to experiencing more events of all kinds. The correlations between events scores and age are quite strong. Younger people experience many more major life events than older people; they are, for example, more likely to get married or divorced, and be promoted or sacked. The

TABLE IV

Links among personality, age and life events

Panel A: Correlations						
E	1.00					
N	−0.18	1.00				
O[a]	0.12*	0.10*	1.00			
Age	−0.20	−0.15	−0.18	1.00		
Pos. Events	0.26	0.12**	0.25	−0.38	1.00	
Neg. Events	0.04[ns]	0.25	0.25	−0.31	−0.47	1.00
Panel B: Regressions for positive and negative life events						
Pos. events = 0.17 E** + 0.22 O***−0.04 Age*** + 1.72**						
$N = 394$, $R^2 = 22.7\%$						
Neg. events = 0.20 N*** + 0.31 O***−0.03 Age*** + 0.79[ns]						
$N = 393$, $R^2 = 16.5\%$						

Notes: O, 'openness to feelings' (Costa and McCrae, 1985).Key: *** Significant at 0.001 level, **significant at 0.01, *significant at 0.05, ns, not significant at 0.05.

regressions in Panel B confirm and summarize these results. In combination E, O and age account for 22.7% of the variance in positive events scores, while N, O and age account for 16.5% of variance in negative events.

Table V compares the life events scores of the four Quadrants. Respondents have been split at the mean on all variables shown.

The evidence in Table V (taken in conjunction with Table IV) shows that Quadrant 1 (the high E, low N group) repeatedly experience many positive events and few negative events. Within Quadrant 1, as within all quadrants, people who score relatively high on O, and also younger people, experience more events of both kinds. Quadrant 2 (low E, high N) experiences many negative events and relatively few positive events. Quadrant 3 is the group to whom 'everything' happens; in each time period they experience more positive events and more negative events than people in the other quadrants. They are the youngest group (mean age = 34) and score highest on O (7.7). Quadrant 4 is the group to whom 'nothing' happens; they experience the fewest positive events, and along with Quadrant 1 also the fewest negative events. Quadrant 4 is the oldest group (mean age = 44) and scores lowest on O (7.3).

Hypothesis 5.1: To the extent that, in any given time period, a person just repeats the pattern of positive and negative life events that is normal for him/her, then WB and IB will remain at or revert to their set points.

Hypothesis 5.2: To the extent that, in a given time period, a person changes/deviates from his/her own normal pattern of events (i.e. relatively exogenous events happen), then WB and/or IB will deviate from their set points.

TABLE V

Panel B: Four quadrants of E and N: Differences in life events scores

	Positive events		Negative events	
	Mean (SD)	Significant level?	Mean (SD)	Significant level?
Quadrant 1 vs. Quadrant 2	4.5 (2.4) 4.1 (2.4)	$p < 0.05$	2.5 (1.4) 3.1 (2.1)	$p < .0l$
Quadrant 1 vs. Quadrant 3	4.5 (2.4) 5.1 (2.5)	ns	2.5 (1.4) 3.3 (2.8)	$p < .001$
Quadrant 1 vs. Quadrant 4	5.1 (2.5) 3.5 (2.0)	$p < 0.001$	2.5 (1.4) 2.5 (1.8)	ns
Quadrant 2 vs. Quadrant 3	4.1 (2.4) 5.1 (2.5)	$p < 0.001$	3.1 (2.1) 3.3 (2.8)	ns
Quadrant 2 vs. Quadrant 4	4.1(2.4) 3.5 (2.0)	$p < 0.05$	3.1 (2.1) 2.5 (1.8)	$p < 0.0l$
Quadrant 3 vs. Quadrant 4	5.1 (2.5)	$p < 0.001$	3.3 (2.8) 2.5 (1.8)	$p < 0.001$

Notes: Quadrant 1 = high E, low N; Quadrant 2 = low E, high N; Quadrant 3 = high E, high N; Quadrant 4 = low E, low N. Quadrant ns: Q1 = 182, Q2 = 157, Q3 = 135, Q4 = 128. ns = not significant at the 0.05 level.

Panel regression fixed effects models are used to test these hypotheses. Prior to testing, it bears repeating that we have earlier shown that within-person levels of WB and IB are quite stable over time. This 'fact' indicates that, in the case of most people, the events which happen to them have only transitory effects, and they soon revert to their set points of WB and IB. If WB and IB were not stable, then the interpretation of the results given below would not hold.[12] The first pair of equations relates to life satisfaction (WB), and the second pair relates to negative affect (IB). In Equations 2.1 and 3.1 all life events are included, whereas in Equations 2.2 and 3.2 only 'objective' events are used.

3.2. *Life Satisfaction*

$$\textbf{Life satis.} = \mathbf{0.08\ Pos.\ events}^{***} - \mathbf{0.08\ Neg.\ events}^{***} + 7.43^{***} \quad (2.1)$$

$N = 2491$, R^2 (within-person) = 3.5%

$$\textbf{Life satis.} = \mathbf{0.14\ Obj.\ Pos.\ events}^{***} - \mathbf{0.11\ Obj.\ neg.\ events}^{***} + 7.50^{***} \quad (2.2)$$

$N = 2482$, R^2 (within-person) = 1.8%

3.3. *Negative Affect*

$$\textbf{Neg. affect} = -\mathbf{0.12}\ \textbf{Pos. events}^{***} + \mathbf{0.16}\ \textbf{Neg. events}^{***} + 2.38^{***} \quad (3.1)$$

$N = 2530,\ R^2$ (within-person) = 3.1%

$$\textbf{Neg. affect} = -\mathbf{0.11}\ \textbf{Obj. pos. events}^{\textbf{ns}} + \mathbf{0.18}\ \textbf{Obj. neg. events}^{***} + 2.35^{***} \quad (3.2)$$

$N = 2520,\ R^2$ (within-person) = 1.0%

Key: ***significant at the 0.001 level; **significant at 0.01; *significant at 0.05; ns, not significant at 0.05.

Recall that these GLS regressions are within-person. It follows that a reasonable interpretation of the significant positive coefficient ($p = 0.001$) for positive life events in Equation 2.1 is: 'In any given short time period (i.e. in any 2-year period; the Australian panel interviewed every 2 years), the more a person changed/deviated upwards from his/her mean positive events scores for the total period (8 years), the higher his/her life satisfaction rose. If, in the short term, positive events fell below their long-term mean, then life satisfaction also fell'. The negative coefficient for negative life events ($p = 0.001$) should be given a parallel interpretation.

Equation 3.1 shows the same pattern in regard to negative affect. Short-term upward and downward changes/deviations from long-term average scores for events were reflected in short term changes in respondents' negative affect scores. Results were significant at the 0.001 level.

Equations 2.2 and 3.2 just confirm that the results hold when the list of events is confined to those which are more or less 'objective'. As it turned out, the effect sizes (bs) were much the same or even a bit larger than in Equations 2.1 and 2.2, although in Equation 3.2 the coefficient for positive events fell short of significance at the 0.05 level.

To recapitulate: there are two key advantages of a fixed effects model in comparison with previously used methods of testing DE theory (Fujita and Diener, 2004; Headey and Wearing, 1989, 1992; Lucas et al., 2003; Magnus et al., 1993). One of these advantages is substantive, the other methodological. The substantive benefit is that a fixed effects model precisely tests the hypothesis that deviations from a person's own mean or set point of happiness are associated with, and presumptively caused by, deviations from their own 'normal' pattern of life events. Previous tests, using OLS, mainly relied on making within-person inferences from across-person analyses (Headey

and Wearing, 1989). The methodological benefit is that a fixed effects model eliminates the potentially biasing influence of unobserved factors.

Hypothesis 6: Major life events are usually anticipated beforehand and habituated to afterwards. These adaptive mechanisms reduce their impact on WB and IB and promote return to set points.

Adaptation theory as formulated by Helson (1964) was based on direct evidence of subjects at first reacting strongly and then habituating to changes in measurable physical conditions (e.g. temperature levels). Subjects' own changing estimates of conditions during the course of experiments demonstrated adaptation. However, use of the term adaptation in SWB research has been based on inference or even tautology. The WB and IB of survey respondents have been observed not to change much, and not for long, in the face of many major life events. So it has been inferred that adaptation has occurred. Specific adaptive mechanisms and the length of time they take to work have not been much researched. Recent work by Diener and colleagues has begun to remedy this (Suh et al., 1996; Lucas et al., 2003; Clark et al., 2004). Using the German panel to follow respondents over time, they have shown that two major events getting married and becoming unemployed – are usually *anticipated* by a year or two. Life satisfaction rises in anticipation of marriage and falls in anticipation of unemployment. After getting married life satisfaction rises further, and after the unemployment axe falls it declines further. It then typically takes a year to 18 months to return to set point, but in some cases *habituation* takes longer, and some people never return to baseline. As noted earlier, repeated spells of unemployment have been shown to have a 'scarring effect', so that many victims never return to baseline (Clark et al., 2004).

Anticipation and habituation can be regarded as adaptive mechanisms which reduce the impact of life events on WB and IB and promote return to set points. It may be noted that, while anticipation is not always possible (i.e. in the case of truly exogenous events), habituation always is. The research by Diener and colleagues represents a major advance, because without explicit and observed adaptive mechanisms DE theory relies too much on inference.

3.4. *Section B: Revising and Extending DE Theory: Person Characteristics and Life Events Associated with Medium Term Changes in SWB*

Accepting the evidence in the first part of the paper, the purpose of this section is to revise DE theory in order to make it more dynamic and

capable of accounting for the finding that, although most people's levels of WB and IB prove to be stable, significant minorities record substantial medium and long-term changes. It should be recognized that this section is more speculative and dependent on *a priori* theorizing than earlier parts.

In fact, it is only becomes feasible to begin to explain longer-term change in SWB because of the very recent availability of personality measures in the long-running German panel (Gerlitz and Schupp, 2005). Personality traits were in fact measured for the first time in this panel in a pilot study conducted in early 2005; the data being made available to the author after the first draft of this paper had been written. The sample size was 8193 and 1102 of these respondents (West Germans only) had recorded their life satisfaction every year since 1985. It is of course possible that the tentative findings here will need to be modified when the main 2005 data set becomes available.[13]

What hypotheses about medium and long-term change can be generated by DE theory? First, to restate the obvious, the large majority of people are predicted *not* to deviate much from their normal (mean) equilibrium levels of WB and IB. Only small minorities are predicted to change. But who is most likely to change?

A reasonable conjecture is that the individuals who are most open to change – in a sense the most high risk people – are those who rate high on E, high on N and high on O. These people are most prone to experience many positive events and many negative events. In a sense they roll the dice more often than other people. So there may be a greater chance that some event or combination of events will occur which have the effect of moving them away from their own normal equilibrium levels of WB and IB.

Extending this line of thinking, we can distinguish between *upside risk* and *downside risk*. High levels of E, especially if combined with low levels of N and high levels of O, may expose a person to upside risk in the sense that many positive events (and rather few negative events) are likely to occur, so there may be a higher than average probability of medium term increases in WB. By contrast, high levels of N, especially if combined with low E and high O, may expose a person to downside risk in that many negative events (and few positive events) are likely to happen, perhaps increasing the chances of medium or long-term increase in IB. Completing the conjecture, we hypothesize that those who rate high on E, N and O will be exposed to high levels of upside and downside risk, and that those who rate low on all three traits will be exposed to low levels of risk in either direction.

Hypothesis 7.1: Even in the medium and long-term most people do not deviate much from their own normal (mean) equilibrium levels of WB and IB. Only small minorities record substantial change.

Hypothesis 7.2: High levels of E and O, combined with low levels of N, are associated with a high upside risk of positive life events and medium term increases in WB.

Hypothesis 7.3: High levels of N and O combined with low levels of E, are associated with a high downside risk of negative life events and medium term increases in IB.

Hypothesis 7.4: High levels of E, N and O are associated with high upside and downside risks in the medium term.

Hypothesis 7.5: Low levels of E, N and O are associated with low upside and downside risks in the medium term.

Clearly these five hypotheses are extensions of the thinking behind the 'quadrants' of WB and IB discussed earlier. They reformulate in dynamic terms ideas which were previously stated as static relationships. However, it should be stressed that the dynamic hypotheses do not follow in any automatic or syllogistic way from the static hypotheses. It could conceivably be the case, for example, that because high E, high O and low N people are more frequently exposed to positive events, they habituate and record less not more change in WB than other people. As usual, it is an empirical issue.

Hypothesis 7.1: evidence that minorities report substantial medium to long-term changes in life satisfaction.

The most straightforward way to assess medium to long-term change in life satisfaction, given 20 years of data, is to select definitions or cut-off points for 'substantial positive change' and 'substantial negative change', and then see what percentage of respondents had substantially higher or lower satisfaction scores in the last few years of the panel than in the first few years. Somewhat arbitrarily, we shall define changes of two or more on the 0–10 life satisfaction scale as 'substantial changes', and then treat change as well established (quasi-permanent?) if a respondent's average rating in the last 5 years of the panel was two or more points higher or lower than in the first 5 years. On this basis, it transpires that 5.5% of the German panel recorded substantial gains in life satisfaction, 13.4% recorded losses, and 81.1% had more or less stable ratings. This supports the hypothesis that, while the large majority stay close to their WB equilibrium levels (or set points), small but not tiny minorities show substantial change.[14]

Hypotheses 7.2–7.5: evidence relating personality traits to medium and long-term change in life satisfaction.

We now attempt to test hypotheses 7.2–7.5 which postulate relationships between personality traits and longer-term changes in SWB. In Equation 4 the dependent variable is the difference between respondents' mean life satisfaction scores in 2000–2004 and their scores in 1985–1989 (Life Satisfaction$_{2000–2004}$–Life Satisfaction$_{1985–1989}$). On the right hand side are the personality traits E[15], N and O, with gender, age, and age squared included as 'controls'. Also included on the RHS is life satisfaction in 1985–1989, because substantial over time regression-to-the-mean in life satisfaction scores invariably has the consequence that first difference scores are negatively related to life satisfaction in the initial time period.

$$\begin{aligned}&\textbf{Life Satis.}_{\textbf{2000–2004}} - \textbf{Life Satis.}_{\textbf{1985–1989}}\\&= +\mathbf{0.07E}^{***} - \mathbf{0.13N}^{***} + \mathbf{0.01\,O}^{\mathrm{ns}} - 0.01\ \mathrm{Female}^{\mathrm{ns}}\\&\quad + 0.02\ \mathrm{Age}^{\mathrm{ns}} - 0.00\ \mathrm{Age}^{2\mathrm{ns}} - 0.56\ \mathrm{Life\ Satis.}^{***}_{(85–89)} + 3.32^{***} \qquad (4)\end{aligned}$$

$N = 1216$, R^2 (between persons) = 28.4%

Key: *** = significant at the 0.001 level; **significant at 0.01; *significant at 0.05; ns, not significant at 0.05.

Our hypotheses relating to E and N are supported, but appear false in relation to O. More extraverted respondents, who also scored low on N (emotionally stable) were more likely than others to record significant increases in life satisfaction during this 20-year period. Individuals who rated high on N and low on E were more likely to record declines in satisfaction. In this equation, personality trait O was unrelated to changes in life satisfaction. The extraversion result could be regarded as in line with research showing that extraverts tend to derive more utility from favourable stimuli than introverts (Larsen and Ketelaar, 1991; Rusting and Larsen, 1997; Lucas and Baird, 2004). Similarly, the result for N could be regarded as broadly in line with laboratory research which indicates that high N individuals amplify the effects of negative stimuli (Larsen, 1992).

The result relating to O is troubling because it is clear from the Australian panel data that O is related to experiencing both more positive life events and more negative events (see Table IV). It could reasonably be argued that Hypotheses 7.2–7.5 imply that interaction terms are required among the personality traits. In particular, it might be predicted that (E * O) would produce increases in life satisfaction, and that (N * O) would produce decreases. In practice, however, two-way and three-way interaction terms added to Equation 4 were either not statistically significant, or, as is often

the case, yielded coefficients with the 'wrong' signs, partly due to multicollinearity. A tentative and interim conclusion is that, while O is related to experiencing both positive and negative life events, it may be unrelated to changes in life satisfaction.

The Australian panel is far from ideal for testing hypotheses about medium term change in SWB. Even so, as Appendix A shows, an equation predicting changes in life satisfaction in Australia between 1981 and 1989 yielded results very similar indeed to those in the German equation. It was also found, in line with our hypotheses, that changes in IB (negative affect) were negatively related to E and positively related to N. They were also strongly negatively related to O. There is some possibility, worth investigating in future research, that O might mainly be related to negative events and changes in IB, rather than to positive events and changes in WB (see Appendix A).[16]

3.5. *Further Evidence Relating Personality to Upside and Downside Risk*

Last, we review evidence about the relationship between personality traits and fluctuations in life satisfaction over time. Results in the previous section showed that high E and high N individuals recorded larger changes in life satisfaction between the starting years and the latest available years of the German panel than respondents with low E and low N ratings. Here we simply confirm that high E and high N scorers recorded larger fluctuations throughout the 20-year period than others. This shows that high E, low N individuals are indeed the people with the highest upside risk of gains in life satisfaction, and that low E, high N people are at greatest downside risk for losses in life satisfaction.

Equation 5 is a random effects GLS equation relating personality to changes in satisfaction in the five year periods 1985–1989, 1990–1994, 1995–1999 and 2000–2004.

$$\begin{aligned}\textbf{Changes in Life Satis.}_{\textbf{1985–2004}} &= \mathbf{0.08\,E}^{***} - \mathbf{0.14\,N}^{***} - \mathbf{0.00\,O}^{\mathbf{ns}} + \mathbf{0.08}\ \text{Female}^{\text{ns}} \\ &\quad - 0.02\,\text{Age}^{\text{ns}} + 0.00\,\text{Age}^{2\text{ns}} + 7.72^{***} \qquad (5)\end{aligned}$$

$N = 5670$, R^2 (between persons) 8.1%

Key: ***Significant at the 0.001 level; **significant at 0.01; *significant at 0.05; ns, not significant at 0.05.

Equation 5 gives very similar results to Equation 4, but it establishes a different point. It shows that the more extreme people's ratings are on E and N, the larger are their fluctuations or deviations from mean levels of

life satisfaction throughout this period. Given the earlier Australian evidence we can take it as read, or at least very likely, that these bigger fluctuations are due to experiencing more rather than fewer life events. High E, high N individuals roll the dice most often: low E, low N people roll it least often.

3.6. *Summary: Profiles by Personality, Life Events and Probability of Medium and Long-term Change in SWB*

We can now summarize by profiling seven patterns of development in SWB over time, indicating which patterns are most common and which are rare (Figure 1).

Profile 1 shows the most common pattern of SWB over time – no change with only quite small fluctuations around an individual's own equilibrium level (mean level) or set point. This individual is near the mean on both E and N and continually experiences life events which are fairly normal or typical for him/her; there are few exogenous shocks.[17] Profiles 2 and 3 can be viewed together and show the next most common patterns; stability of SWB over time but with quite wide short term fluctuations. Individual 2 is a quite a 'happy' person, being high on E and middling or low on N. He/she continually experiences quite a lot of favourable events and few negative events. Most of these events are of a kind which are normal for him/her (given the personality profile), but because of 'high risk' some exogenous events occur, and fluctuations around his/her mean SWB are quite large. Profile 3 is of an individual who is quite 'unhappy', being middling or low on E and high on N. Events are mainly negative and fluctuations around his/her mean SWB are quite large.

Profiles 4 and 5 are of individuals who are about equally 'happy' and are close to the mean for the population, even though they have very different personalities. Individual 4 is high on E and on N. Lots of favourable and negative events continually occur, and wide fluctuations occur around his/her mean SWB. Individual 5 rates low on both E and N, few events occur and fluctuations in SWB are small.

Profiles 6 and 7 are of individuals representing small minorities. Profile 6 is a high E, low N person who is shown to have experienced a sharp rise in SWB. For a while, at the beginning of the period, this person experienced events which were normal for him/her. But then some partly exogenous positive event or combination of events happened and produced a sharp rise. Although this event(s) was somewhat unpredictable, the fact that the

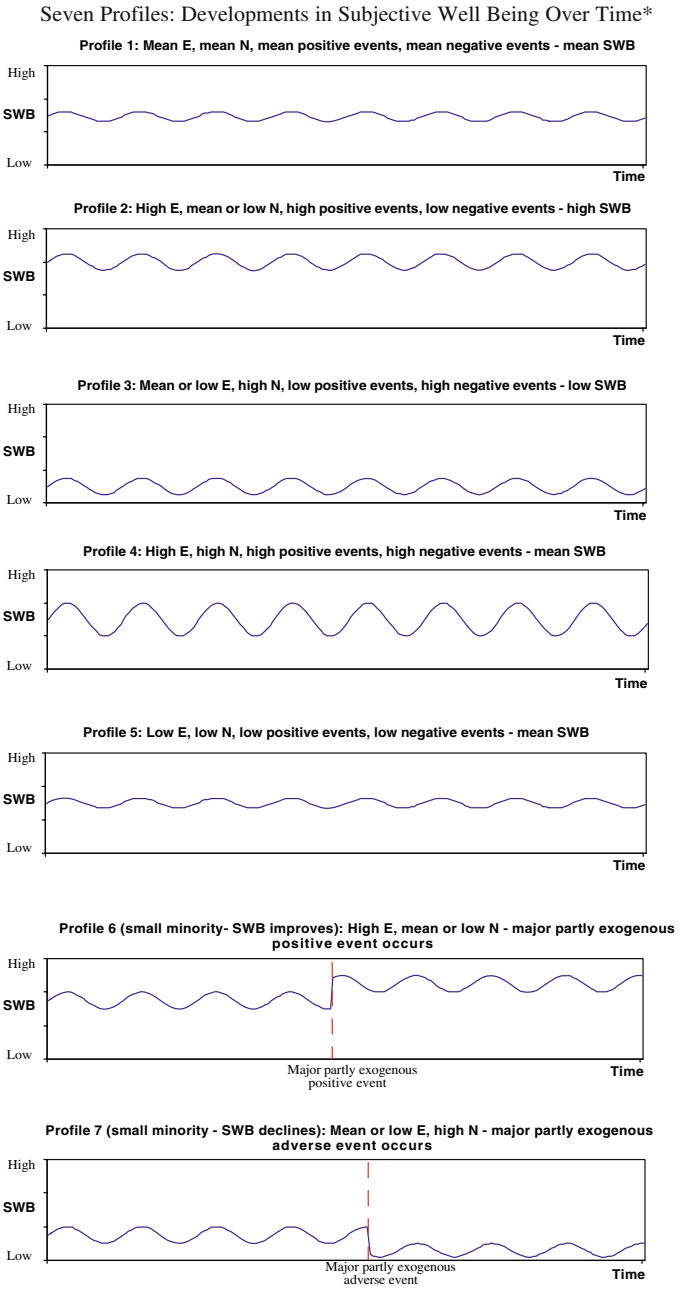

Fig. 1. Seven profiles: Developments in subjective well being over time. SWB scores can be thought of here as standardized (mean of 0, standard deviation of 1.0).

person was high E, low N made it not all that unlikely. Finally, Profile 7 shows a developmental pattern for a low E, high N person to whom a major partly exogenous negative event happens.

It is important to say which logically possible patterns of SWB development are predicted by DE theory to be very *un*likely. It is unlikely that a person who scores near the mean on E and N will experience major exogenous events of a drastic enough kind to move him/her permanently away from his/her equilibrium levels of WB and IB. Of course, it could happen. As previously noted, it has been shown that some events, most notably the death of one's child, are so dreadful as to permanently lower SWB, and obviously that could happen to any parent. Another prediction that follows from DE theory is that a one-off upward or downward change in SWB is more likely than (say) 10 or 20 years of continuous gain or decline. This is simply because of the improbability of major events of a kind which are abnormal for a particular person repeatedly happening to him/her.

4. DISCUSSION

The purpose of this paper has been to reassess and revise the DE theory of SWB, employing longer-term data and more powerful methods of panel analysis than previously available. Prior research had shown that personality traits account for repeating patterns of life events, as well as for equilibrium levels (or set points) of WB and IB (Headey and Wearing, 1989). The further claim that people only move away from their set points of WB and IB if the events which happen to them deviate from their own 'normal' pattern was not well established, although it seemed a logical inference. In this paper the claim has been supported by *within person* panel regression results which precisely relate two-year changes in WB and IB to simultaneous deviations-from-the-mean (i.e. from each person's own mean) in life events scores. A second contribution of the paper was to try and extend DE theory to account for an important *between-person* difference? Why is it that, while most people record stable levels of WB and IB, some people record substantial medium and long-term changes? *Between-person* panel regression analysis, based on 'best-available' but not ideal data, gave tentative support to hypotheses about the personality traits and life events experienced by people most prone to upside or downside risk; that is, risk of medium tern gains or losses in SWB.

It may be useful to conclude with some comparisons between DE theory and alternative theories of SWB. DE theory, it is suggested, has the merit of accounting (or partly accounting) for a wider range of results than other

widely cited theories. Its strength is that it integrates results relating to personality, life events, WB and IB. Competing theories of SWB include adaptation theory and its offshoots (Helson, 1964; Brickman and Campbell, 1971; Brickman et al., 1978), theories related to personality and heritability (Costa and McCrae, 1980ab; Costa, McCrae and Zonderman, 1987; Lykken and Tellegen, 1996; Lykken, 1999), social comparison theory and offshoots including multiple discrepancies theory (Easterlin, 1974; Michalos, 1985), and goal-setting or telic theory (Emmons, 1986; Cantor and Sanderson, 1999; Peterson, 1999; and for reviews see Diener et al., 1999; Diener and Lucas, 1999; Veenhoven, 1994).[18] Adaptation theory offers an understanding of linkages between life experiences and SWB, but says nothing about personality. Theories based on personality, and on heritability more generally, say nothing about life events. Social comparison theory and its relatives, including multiple discrepancies theory, extend our knowledge of the range of yardsticks against which people assess their current life and arrive at their SWB ratings, but say nothing about personality or life events. Goal setting or telic theories tell us about linkages between SWB, the different types of goals people pursue, and their successes and failures in pursuing these goals. Again the scope is quite limited.

If DE theory has value, it will be important to try and keep improving its specification and extending its scope, so that it does a better job of accounting for change. This paper has only made small beginnings in accounting for medium to long-term changes in SWB. It may be that in order to understand such changes, we need to conceive of slow changes in personality accompanied by slow changes in individuals' patterns of life events. Such changes might best be modeled using moving average specifications. Such models, sometimes labeled ARMA models, are available, but they would require better quality long-term data than we have at present. Inclusion of personality measures and reasonably detailed life events inventories in existing national panel studies are one way forward. It would also be ideal if the self-report measures of SWB in these surveys could be supplemented by high quality 'on line' data.

ACKNOWLEDGEMENTS

Many thanks to Alex Wearing of the University of Melbourne for comments on an earlier draft. Also thanks to Gert Wagner, Director of the German (SOEP) panel, who made available 2005 pilot study data on personality and well-being. Valuable comments were also received from an anonymous SIR reviewer.

APPENDIX A

The Australian Panel: Covariates of Change in WB and IB (OLS Regressions)

	Life satis.$_{89}$–Life satis.$_{81}$	Neg. affect$_{89}$–Neg. affect$_{81}$
Extraversion	0.14*	−0.26*
Neuroticism	−0.19***	0.31***
Openness	−0.05^{ns}	0.51***
Female (1–0)	−0.29^{ns}	−0.35
Age	−0.09*	−0.09 ns
Age^2 (/1.0)	0.01*	0.01 ns
Life satisfaction$_{81}$/ Negative affect$_{81}$	−0.58***	−0.60***
Constant	5.89***	0.91 ns
N	283	283
R^2	32.2%	28.2%

***Significant at 0.001 level; **Significant at 0.01; *Significant at 0.05; ns, not significant at 0.05.

NOTES

[1] However, some researchers propose additional sub-dimensions. WB can be split into life satisfaction and positive affect. IB can be split into anxiety and depression (and other less common adverse states).

[2] Unlike E and N, the personality trait O – Openness to Experience – appears not to be directly linked to well-being and ill-being. Its relevance lies in being positively correlated with the frequency of experiencing both positive and negative life events.

[3] As noted in Section B, data for a 2005 pilot study, which included measures of personality, became available after this article had been drafted. Results relating to the pilot study have been included to Section B.

[4] A life satisfaction item was included in 1996–2000 and then dropped. The life satisfaction data are not used in this paper.

[5] Econometricians vary in the advice they give on these issues Baltagi (1995) takes the view that a random effects model is usually appropriate in analysis of household panels, even if assumptions relating to errors are not met. In his view a fixed effects model should only be used if one has specific hypotheses about within-person fixed effects which have not been measured and which are likely to affect the dependent variable.

[6] In the regression analyses reported below only a 'balanced' panel of respondents who participated in all survey waves was used.

[7] Due to the small (negative) correlation between E and N, Quadrants 3 and 4 have fewer cases than Quadrants 1 and 2.

[8] The positive events treated as 'objective' were: you were promoted; you passed an important exam; you became engaged; you got married; husband and wife got together again after separation; you experienced a religious conversion. The negative events treated as 'objective' were: you were unemployed; sacked or laid off; you had a major financial crisis; your own business failed; you failed an important exam; you had a serious illness or injury; you had a

serious accident; you broke off an engagement; you separated from your spouse; you divorced; your spouse died; a child of yours died; another close family member died; a close friend died; you (your wife) had a miscarriage, abortion or still birth; you were robbed; you were physically assaulted; problems with the police leading to a court appearance; prison sentence; you had a civil suit (e.g. divorce, custody, debt).

[9] This is true in all cells where the entries cover the complete list of events. In two of the 16 cells where only objective events are included, correlations are close to zero.

[10] In an earlier paper (Headey and Wearing, 1989), we showed that some specific events kept happening to the same people. These included positive friendship and job related events, and negative financial and job related events.

[11] It is not suggested that specific events can be classified as either exogenous or endogenous. Almost any particular event could be partly endogenously caused (brought on by the characteristics of the person to whom it happened), and partly exogenous (due to outside forces).

[12] If the within-person means changed over time, a possible interpretation would be that life events were driving change.

[13] Data from the main study are expected in 2006.

[14] The apparently larger number of downward changes, compared to upward changes, is probably to ceiling effects. That is, people who are, in the early years of the panel, at the top of the 0–10 scale cannot go higher. There are many more such people than those who start at the bottom end of the scale. Also, there appears to be some tendency for people to over-state their life satisfaction in the first years of a panel, compared to later years, presumably for social desirability reasons (see Frijters et al., 2004).

[15] A modification was made to the measure of E. One self-description item, 'reserved', had the effect of much reducing the correlation between E and life satisfaction; in fact making it lower than the correlation between O and life satisfaction. This outcome was so out of line with previous research (e.g. Costa and McCrae, 1991) that it seemed better to drop the item.

[16] Again, interaction terms were either statistically non-significant or substantively uninterpretable.

[17] The profiles could be elaborated by including O and age, but the differences would only be of degree; in statistical terms a little more variance would be accounted for in life events and in WB and IB.

[18] This list only includes theories built mainly on global self-reports of SWB. It does not include theories which rely on detailed 'on-line' evidence of cognitive and affective processing (see esp. Parducci, 1995; Schwarz and Strack, 1999; Kahneman, 1999). These latter theories offer new insights into the shape of WB/utility curves; insights not accessible from global data.

REFERENCES

Andrews, F.M. and S.B. Withey: 1976, Social Indicators of Well-Being (Plenum New York).

Argyle, M.: 1987, The Psychology of Happiness (Methuen, Oxford).

Baltagi, B.H.: 1995, Econometric Analysis of Pond Data (John Wiley, New York).

Becker, G.: 1975, Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education (National Bureau of Educational Research New York).

Bradburn, N.M.: 1969, The Structure of Psychological Well-Being (Aldine, Chicago).

Breusch, T.S. and A.R. Pagan: 1980, 'The Lagrange multiplier test and its applications to model specification in econometrics', Review of Economic Studies 12, 33–56.

Brickman, P.D. and D.T. Campbell: 1971, 'Hedonic relativism and planning the good society', in M.H. Appleby (ed.), Adaptation-Level Theory (Academic Press, New York), pp. 287–302.

Brickman, P.D., D. Coates and Janoff-Bulman: 1978, 'Lottery winners and accident victims: Is happiness relative?', Journal of Personality and Social Psychology 36, 917–27.

Cantor, N. and C.A. Sanderson: 1999, 'Life task participation and well-being: The importance of taking part in daily life', in D. Kahneman, E. Diener and N. Schwarz (eds.), Well-Being: The Foundations of Hedonic Psychology (Russell Sage, New York), pp. 230–243.
Clark, A.E., Y. Georgellis, R.E. Lucas and E. Diener: 2004, 'Unemployment alters the set point for life satisfaction', Psychological Science 15, 8–13.
Costa, P.T. and R.R. McCrae: 1980a, 'Influences of extraversion and neuroticism on subjective well-being', Journal of Personality and Social Psychology 38, 668–678.
Costa, P.T. and R.R. McCrae: 1980b, 'Still stable after all these years: Personality as a key to some issues in adulthood and old age', in P.B. Baltes and O.G. Brim (eds.), Life Span Development and Behavior (Vol. 3) (Academic Press New York), pp. 65–82.
Costa, P.T. and R.R. McCrae: 1985, The NEO Personality Inventory (PAR, Odessa, Fla).
Costa, P.T. and R.R. McCrae: 1991, NEO PI-R (PAR, Odessa, Fla).
Costa, P.T., R.R. McCrae and A.B. Zonderman: 1987, 'Environmental and dispositional influences on well-being. Longitudinal follow-up of an American national sample', British Journal of Psychology 78, 299–306.
Diener, E., E.M. Suh, R.E. Lucas and H.L. Smith: 1999, 'Subjective well-being: Three decades of progress', Psychological Bulletin 25, 276–302.
Diener, E. and R.E. Lucas: 1999, in D. Kahneman, E. Diener, N. Schwarz (eds.), Well-being: The Foundations of Hedonic Psychology (Russell Sage, New York), pp. 213–229.
Diener, E., R.A. Emmons, R.J. Larsen and S. Griffen: 1985, 'The satisfaction with life scale', Journal of Personality Assessment 49, 71–75.
Easterlin, R.A.: 1974, 'Does economic growth improve the human lot? Some empirical evidence', in P.A. David and M.W. Reder (eds.), Nations and Households in Economic Growth (Academic Press, New York), pp. 89–125.
Emmons, R.A.: 1986, 'Personal strivings: An approach to personality and subjective well-being', Journal of Personality and Social Psychology 51, 1058–1068.
Eysenck, H.J. and S.B.G. Eysenck: 1964, Personality Structure and Measurement (Routledge & Kegan Paul, London).
Frederick, S. and G. Loewenstein: 1999, in D. Kahneman, E. Diener and N. Schwarz (eds.), Well-Being: The Foundations of Hedonic Psychology (Russell Sage, New York), pp. 302–329.
Frijters, P., J.P. Haisken-Denew and M.A. Shields: 2004, 'Money does matter! Evidence from increasing real incomes and life satisfaction in East Germany following reunification', American Economic Review 94, 730–741.
Fujita, F. and Diener: 2005, 'Life satisfaction set-point: Stability and change', Journal of Personality and Social Psycholgy 88, 158–164.
Gerlitz, J.-Y. and J. Schupp: 2005, 'Zur Erhebung der Big-Five-basierten Persoenlichkeitsmerkmale im SOEP', www.diw.de/deutsche/produkte/publikationen/researchnotes/docs/papers/rn4.pdf.
Goldberg, D.: 1978, Manual of the General Health Questionnaire (NFER, Windsor).
Hausman, J.A.: 1978, 'Specification tests in econometrics', Econometrica 46, 1251–1271.
Headey, B.W. and A.J. Wearing: 1989, 'Personality, life events and subjective well-being: Toward a dynamic equilibrium model', Journal of Personality and Social Psychology 57, 731–739.
Headey, B.W., J. Kelley and A.J. Wearing: 1993, 'Dimensions of mental health: life satisfaction, positive affect, anxiety and depression', Social Indicators Research 29, 29–63.
Headey, B.W. and A.J. Wearing: 1992, Understanding Happiness: A Theory of Subjective Well-being (Longman Cheshire, Melbourne).
Helson, H.: 1964, Adaptation-Level Theory (Harper & Row, New York).
Henderson, S., D.G. Byrnes and P. Duncan-Jones: 1981, Neurosis and the Social Environment (Academic Press, New York).

Holmes, T.H. and R.H. Rahe: 1967, 'The social readjustment rating scale', Journal of Psychosomatic Research 11, 389–395.
Kahneman, D.: 1999, 'Objective happiness', in D. Kahneman, E. Diener and N. Schwarz (eds.), Well-Being: The Foundations of Hedonic Psychology (Russell Sage, New York), pp. 3–25.
Krause, P.: 2003, 'Income, poverty and dynamics in Germany', in P. Krause, G. Baecker and W. Haensch (eds.), Combating Poverty in Europe (Ashgate, Aldershot), pp. 93–116.
Larsen, R.J.: 1992, 'Neuroticism and selective encoding and recall of symptoms: Evidence from a combined concurrent-retrospective study', Journal of Personality and Social Psychology 62, 489–498.
Larsen, R.J. and T. Ketelaar: 1991, 'Personality and susceptibility to positive and negative emotional states', Journal of Personality and Social Psychology 61, 132–140.
Larsen, R.J., E. Diener and Emmons: 1985, 'An evaluation of subjective well-being measures', Social Indicators Research 17, 1–18.
Lillard, L.A. and R.J. Willis: 1978, 'Dynamic aspects of earnings mobility', Econometrica 46, 985–1012.
Lucas, R.E., A.B. Clark, Y. Georgellis and E. Diener: 2003, 'Reexamining adaptation and the set point model of happiness: Reactions to change in marital status', Journal of Personality and Social Psychology 84, 527–539.
Lucas, R.E. and B.M. Baird: 2004, 'Extraversion and emotional reactivity', Journal of Personality and Social Psychology 86, 473–485.
Lykken, D. and A. Tellegen: 1996, 'Happiness is a stochastic phenomenon', Psychological Science 7, 186–189.
Lykken, D.: 1999, Happiness: What Studies on Twins Show Us about Nature, Nurture and The Happiness Set Point (Golden Books, New York).
Maddi, S.R., P.T. Bartone and M.C. Puccetti: 1987, 'Stressful events are indeed a factor in physical illness: Reply to Schroeder and Costa (1984)', Journal of Personality and Social Psycholology 52, 833–843.
Magnus, K., E. Diener, F. Fujita and W. Pavot: 1993, 'Extraversion and neuroticism as predictors of objective life events: A longitudinal analysis', Journal of Personality and Social Psychology 65, 1046–1053.
Michalos, A.C.: 1985, 'Multiple discrepancies theory (MDT)', Social Indicators Research 16, 347–413.
Parducci, A.: 1995, Happiness, Pleasure and Judgment: The Contextual Theory and its Applications (Erlbaum, Hove, England).
Peterson, C.: 1999, 'Personal control and well-being', in D. Kahneman, E. Diener and N. Schwarz (eds.), Well-Being: The Foundations of Hedonic Psychology (Russell Sage, New York), pp. 288–301.
Rahe, R.H. and R.H. Arthur: 1978, 'Life change and illness studies', Journal of Human Stress 4, 3–15.
Rowlison, R.T. and R.D. Feiner: 1988, 'Major life events, hassles, and adaptation in adolescence: Confounding in the conceptualization and measurement of life stress and adjustment revisited', Journal of Personality and Social Psychology 55, 853–863.
Rusting, C.L. and R.J. Larsen: 1997, 'Extraversion, neuroticism and susceptibility to positive and negative affect', Personality and Individual Differences 22, 607–612.
Schroeder, D.H. and P.T Costa: 1984, 'Influence of life event stress on physical illness: Substantive effects or methodological flaws?', Journal of Personality and Social Psychology 46, 853–863.
Schwarz, N. and F. Strack: 1999, 'Reports of subjective well-being: Judgmental processes and their methodological implications', in D. Kahneman, E. Diener and N Schwarz (eds.), Well-Being: The Foundations of Hedonic Psychology (Russell Sage, New York), pp. 61–84.

Suh, E.N., E. Diener and F. Fujita: 1996, 'Events and subjective well-being: Only recent events matter', Journal of Personality and Social Psychology 70, 1091–1102.
Veenhoven, R.: 1994, 'Correlates of Happiness; 7836 Findings from 603 Studies in 69 Nations: 1911–1994' (Unpublished manuscript, Erasmus University Rotterdam, The Netherlands).
Watson, D., L.A. Clark and A. Tellegen: 1988, 'Development and validation of brief measures of positive and negative affect: The PANAS scales', Journal of Personality and Social Psychology 54, 1063–1070.
Wengle, H.: 1986, 'The psychology of cosmetic surgery: A critical overview of the literature 1960–1982. Part 1', Annals of Plastic Surgery 16, 435–443.
Wilson, W.: 1967, 'Correlates of avowed happiness', Psychological Bulletin 67, 294–306.
Wortman, C.B. and R.C. Silver: 1987, 'Coping with irrevocable loss', in G.R. Vanderbos and B.K. Bryant (eds.), Cataclysms, Crises, Catastrophes: Psychology in Action (APA, Washington, DC).

Melbourne Institute B. Headey
University of Melbourne
Parkville, Australia 3052