In the tradition of North American psychological research on stress, paper and pencil measurements have been long-time favourites. Given problems of validity and reliability of checklists and self-reports, Lazarus and Folkman's book in 1984 *Stress, Appraisal and Coping* had come to many as a consolation for the difficulty of measuring life hardships in a consistent fashion: the quintessence of the stress experience lies in the subjective appraisal of threat.

Certainly the work by Lazarus and Folkman has been a landmark in the field of psychological stress in America. It inspired a great number of research studies and led the growing attention given to the role of cognition and emotion in mental health (Cronkite and Moos, 1984; Kanner et al., 1981; Vinokur and Caplan, 1986). It also brought an answer to some shortcomings and crises of life event checklists (Tausig, 1982; Zimmerman, 1983). It offered an alternative to the wide use of either counting raw numbers of ticked items or adding up generalised normative change weights (Holmes and Rahe, 1967). However, the popularity of the appraisal approach to stress engulfs much of the research in contamination and circularity (Delongis et al., 1982; Dohrenwend et al., 1984; Hobfoll, Schwarzer and Koo, 1996; Ensel and Lin, 1991; Monroe, 1982).

In physics, it is common practice to describe movements with respect to relative frames of references, especially so in the context of multiple motions. It has proven crucial to our understanding of impact. How well can we appreciate the speed of a train from the word of the traveller walking in one of its coaches? A better description of the situation is achieved by adding the reference frame of the engineer on the tracks or that of a bystander on the platform. Similarly, the dynamics of mental health, life-events, behaviours and cognition may also benefit from a multi-referential description.

Influenced by Kurt Lewin's (1951) Person-Environment formula, for my understanding of the stress process I needed an external frame of reference against which I could appreciate people's appraisal of their life situation. I was interested in what conditions stressed people more than others. How did emotion, perception and cognition change with the course of events? Was it people's view that changed or the circumstances they were in (Bandura, 1997)? Did people 'distort' as most cognitivists put it, have 'negative views' or 'positive illusions' (Beck, 1976; Hammen and Mayol, 1982; Sarason, 1975; Taylor, 1983)? What part was 'reality'
playing? How accurate was people’s appraisal? How reliable? How consensual? The issue was not to find an ‘absolute stressing metric’ but to disentangle subjective appraisal from factual evidence so that some triangulation one could better appreciate the role of individual differences, without necessarily evoking mental illness (Lemyre, 1987).

In my quest for another referencing system, the work of George Brown and his colleagues provided me with a powerful tool for many of my queries. Social Origins of Depression (Brown and Harris, 1978) came as a revelation. Although my own interest lay more in the stress process in the normal population, I had found a meaningful psychosocial thesis of pathology. It was a comprehensive, integrative, articulated and powerful view of the forces in one’s life story: the rarest and finest combination of psychology, sociology, psychiatry and epidemiology. I had also found an accessible, feasible, realistic methodology. Maybe it was not perfect, maybe it was not 100 per cent accurate, but it was consistent and independent.

After my post-doctoral training with the Bedford Square unit, I started to use the Life Event and Difficulty Schedule (LEDS) and its derived complements (for social support and coping) in my work on the cognitive appraisal process in stress, relating contextual and subjective evaluations of life circumstances. Appraisal became my dependent variable and the LEDS contextual ratings the anchor in assessing the environment input.

Subjective appraisal of joint events in couples

The question of objective differences in appraisal has been difficult to document. How much of the variance is due to variation in the stressors being examined versus idiosyncrasies of either perception or reporting? Colette Bron (1993), one of my doctoral students, and I used the LEDS with couples consulting at fertility clinics. We relied on the interviews done independently with each spouse to:

- identify stressors common to both partners
- have their version of each stressor
- get their respective subjective appraisal in a self-reported format on scales matching in content the main ratings of the LEDS.

All interviews were rated blind to the other spouse version. We then identified stressors which were contextually rated as ‘joint’ events that should have been reported by both partners. Data supported the validity of the LEDS: 84 per cent of those rated as joint events were indeed reported by both partners, more so for ‘severe events’.

We compared subjective appraisal ratings of the two partners for joint events that had been given the same contextual severity. (In some instances, joint events were given different severity ratings for each spouse.) We also analysed the differences in terms of whether those joint events were fertility-related or not. Appraisal was measured using a ten-item subjective appraisal rating scale yielding three main numerical components: perceived impact (relating to threat, negativity, loss, severity, undesirability); perceived mastery (relating to control, positivity, challenge); and perceived uncertainty (challenge, danger, unknown, lack of control).

A significant main effect showed that indeed severe events were rated with more impact than non-severe events. Moreover, there was a gender effect. Women reported less stress and more perceived uncertainty than their partners about both fertility and non-fertility events of the same contextual threat. There was also an interaction effect: discrepancies between genders were stronger for chronic difficulties than for events. Women reported higher impact for chronic difficulties, even for contextually mild difficulties, and much more so than their partners. Men reported on the whole that chronic difficulties had little impact. Using multiple regression we found women’s distress was related to the interaction between contextual severity and their appraisal of impact and of uncertainty. The only significant predictor for male distress was appraisal of mastery, independently of contextual severity.

The LEDS provided us with a unique tool to compare responses of spouses to a common situation. The contextual approach ensured that where ratings were comparable, there was an objective and replicable standard that could be related to their subjective responses. As just noted, the subjective impact of events blindly rated as objectively non-severe, was greater for women and much more so for chronic difficulties at all levels of severity. There was also a main effect for gender on appraisal of mastery and a trend towards a similar effect for appraisal of uncertainty. It also looked as though appraisal played a different role in non-severe circumstances than in severe conditions. Thus investigation of cognition, distortion or illusion should not presume a linear relationship. Only the rating system of the LEDS, where contextual and subjective threat/pleasantness are distinguished, could have provided us with such discriminating results.

Naczko, Edwards and Brown (1997) carried out a somewhat similar study of couples especially selected because they had experienced a common severe life event. Instead of emphasizing differences in subjective response, they focused on the development of clinically relevant depression, and found the marked excess of depression among women was due to their greater sensitivity to severe events involving children, preoccupation and housing. Rates of onset following all other types of severe event were comparable between the genders. In both studies the LEDS therefore provided a scaffold from which gender differences in response could be fruitfully explored.

Subjective and contextual evaluations in anticipation of events

The longitudinal design of the infertility study also provided us with an opportunity to investigate anticipated events. Life event research has concentrated on the impact of events occurring in a defined period, mostly in the past six, twelve or more months (Dohrenwend and Dohrenwend, 1974; Kessler et al., 1991; Thoits, 1982). But on occasion events that are anticipated may also be of relevance. Some work on anticipation has been done by Frankenhaus (1980), Antonovsky...
create investigator-based ratings for situations that might be expected to be anticipated, given the context. For example, a housing move for a couple in fertility treatment living in a studio, or a financial difficulty if the woman was the only bread-earner with no work compensation benefits.

b) include events that were subjectively anticipated by the respondent irrespective of the degree to which the current circumstances justified this, and rate them on contextual severity with a special code. Along with these was an estimate of the likelihood of occurrence in the next year.

Although this exercise was unprecedented, agreement between partners interviewed independently about such reported anticipated future events was 67 per cent. Inter-rater agreement on severity of threat and the other major scales yielded a value of 'kappa' similar to that for events that had actually occurred. On a test/retest sub-sample, 90 per cent of anticipated severe events were repeated. The expanded LEDS thus appeared to provide reliable means of gathering data on anticipated events.

Overall, the events anticipated reflected those that had already occurred. Proportions of events were similar across domains for past events and future events except for the health domain. Although in terms of the past twelve months health events had been the most frequently reported, they were the least anticipated for the next twelve months. The overall number of anticipated future events was correlated to the number reported in the prior twelve months. However, proportionally more positive events were anticipated than had occurred in the past twelve months. Women anticipated far more events than men. Finally, the number of events anticipated related to feelings of current stress even when the contribution of recently occurring events was controlled.

This line of work, although at present somewhat tangential to the main thrust of research, attests in its own special way to the power and flexibility of the LEDS system. Over the last twenty or so years, it has become clear that one of its major assets is the way in which new scales can be developed and – it might be added – encouraged by the richness of the data collected by routine LEDS usage. In my experience, there is no sense of being locked into an unalterable instrument; and where a new problem is being tackled it is an ideal platform from which to develop insights and new scales (Castello and Devins, 1988). Furthermore, the conceptualisation of the contextual frame of reference permits a systematic derivation for each psychosocial variable of its mirror images, first from the respondent's perspective and second, from the observer's perspective (i.e. the investigator).

Figure and background changes: peripheral versus focal stressor

The LEDS also provided us with a window on to the appraisal process that takes place in the context of an unfolding event. One study allowed us to gain some insights on how appraisal of one stressor depends on the co-existence of other stressors. With Lisa Fillion (1993), women who were going through a biopsy for suspected breast cancer were interviewed before surgery, after diagnosis and at about six months follow-up, about their life as a whole. One of our interests was how far their appraisal of LEDS-type events and difficulties would change in time and be modified by the tumour's status. Women described in a LEDS format interview the discovery of the suspicious lump in their breast as well as their other events and difficulties. These were rated according to the LEDS guidelines by a team of trained assistants. Women also filled in at each study point a subjective appraisal rating scale relative to:

a) their tumour
b) their worst other stressor.

Thirty women with a malignant tumour were matched on age, income, education and marital status with one of sixty with a benign mass.

Among the striking results was the fact that levels of reported stress were lower at follow-up for the cancer group than the benign group. Moreover, although for the great majority of women the contextual (i.e. objective) severity of their worst other stressor – called here the peripheral stressor – did not change through the study period, their appraisal of this continuing difficulty did change – generally in the direction of lesser stress. There was also an interaction effect associated with the tumour's status. For the benign group, the peripheral stressor lost perceived importance during the diagnosis period but regained its level of preoccupation after the biopsy results. For the malignant group, the peripheral stressor remained at a lower level.

These results speak of the relativity of the appraisal process. They probably accurately indicate how cognitions change according to the context but they also, on a more pessimistic note, underline their unreliability as a metric when a person is living through a developing crisis.

Variation in appraisal for a constant context

Similarly, with Lisa Sweet (1999) we used the LEDS in the earlier stage of breast cancer screening with women age 50 and more, without specific risk factors (no personal nor family history of cancer), participating in a mass prevention programme. Among 800 women who filled appraisal ratings at three different points in time (T1: prior to screening; T2: awaiting results; T3: post results). Three months later, 120 were seen at a follow-up interview (T4). Half of these had received within days of their screening a false positive result based on either their
clinical palpation or mammogram. The other half had received a perfectly negative screen result. All were actually cleared from any detectable malignancy within a few weeks and were known to be healthy at time of interview.

Despite the fact that those with a negative screen result were at the same level of contextually severe throughout the screen process, their appraisals varied a good deal - both between the women and across time. As expected, subjective severity tended to diminish with time and there was for many women a positive change even before the screen results were known. Again, appraisal appeared to be a highly labile process and to use it as the sole indicator of the stressfulness of a situation, as is commonly done, may, on the basis of these findings, be misleading. Our results also call attention to some paradoxical effects of the screening procedure for breast cancer in the sense that in some instances it inappropriately and prematurely reassured women on their personal health while it led to over-estimated population risk.

'I know what I do' or 'I know what to do': rating reported coping

In the same study with Joanne Savoie (1999), we looked at coping using a LEDS-type technique of rating coping. Congruent with a meta-analysis of the literature (Carver and Scheier, 1994; Ender and Parker, 1990; Pearlman, 1991), four coping scales were rated, based on accounts of behaviours and plans: cognitive approach, cognitive avoidance, behavioural approach and behavioural avoidance. Compared to the moderately low reliability of self-reported coping questionnaires, this measure (the LEDS-Cope) obtained good stability. Inter-rater agreement was around .90, test/test was very high at .95. However, the concordance between the respondent's self-rating and the investigator's rating of the person's coping was very low, correlating only .30. Women reported more active coping than judged by investigator-based ratings. Conversely, they reported less avoidant coping than was estimated. Analyses comparing groups and times revealed significant differences only on the contextual measures and not on subjective appraisal. For example, women who had been screened as positive were rated with less behavioural approach coping at T2 and T3 and with more avoidance post-results. Self-reported coping strategies did not differ in a consistent fashion between times or across screen results nor predict distress.

To our knowledge this is one of the first studies contrasting a self-reported measure of coping with an investigator-based one. It proved very helpful in pinpointing the subtleties of the construct of coping and the difficulties in its measurement. Coping scales have been known to show weaknesses but few studies have actually targeted those deficiencies. The differences between self-reports and investigator-based ratings are a matter of hypotheses themselves. It could be that people are unreliable reporters, bad observers of themselves, too susceptible to the effects of social desirability, biased towards the positive illusion of being a good copier, or just telling us that they know what they should do almost in an effort to actually induce in themselves the desirable answer. It is yet to be further explored. The investigator-based system of Brown and colleagues is a golden tool for doing it.

The use of investigator-based measures such as the LEDS

The fundamental characteristic of the LEDS is that it relies on ratings independent of the respondent. The investigators make the judgments. The power of the approach is its flexibility and ability to tackle a surprising range of difficult questions. Certainly, since it still relies on information provided by the respondent, it is open to potential bias. This is true. But to the extent that the questioning attempts to reconstruct actual coping and other behaviour in the course of discussing the development of the crisis, it may well provide a reasonable approximation of what occurred. Direct observations would be preferable but are rarely possible. The flexibility of the LEDS also allows the derivation of new scales custom-made for particular objectives, and the procedure to be followed in developing new scales is now well rehearsed and documented (see chapters 1 and 5 in this volume). One only needs to define the scales and establish their reliability.

But I should be honest. There is a love-hate relationship with the LEDS-system. It costs so much in time, effort, details and agonies. Yet once you have found this level of refinement of information, you cannot go back to bulk undifferentiated data. I cannot live without it.

For a Galilean psychology

In conclusion, I would like to advocate an understanding of psychological processes with more built-in relativity. Especially for psychologists who tend to put the burden of mental health on the psyche of the person, it is good to be more aware and reminded of the weight and inertia of life adversities. Contextual severity has a main effect. Subjective appraisal may add an interaction effect. We should also convey how mental health is a two-way road. Once distress has set in the course of appraisal is modified. Following Copernicus, I would like to challenge the common belief that the Self is the centre of the psychological universe. I rather conceive the Self in orbit around life circumstances whose severity determines the gravitational force. Our role might be to describe constellations that allow comets to change their course.

In this perspective, the etiological model elaborated by Brown and colleagues which articulates a structural chain with an affective chain to describe a final common pathway leading to depression (Akiskal and McKinney, 1973) is a sophisticated integration of sociological factors, psychological processes and psychiatric states. Like Galileo's telescope, the LEDS is the tool to disentangle the various forces involved in the system and help identify the various components. To understand the complexity of the galaxy of life experiences we need more empirical observations about the sequencing of events, emotions, behaviours and cognition. It calls for a double axis paradigm: subjective and contextual.
References


