



Editorial

The genetic tidal wave finally reached our shores: Will it be the catalyst for a critical overhaul of the way we think and do science?

As an undergraduate student in the 1980s, I first had the chance to read Marvin Dunnette's classic article in the *American Psychologist*, entitled "Fads, fashions, and folderol in psychology" (Dunnette, 1966). I don't think I understood much during that first reading but I have since re-read the article numerous times, each time finding something new to ponder and each time being impressed by the author's critical and timeless insights. Therein, Dunnette eloquently describes various "games" that psychologists like to play. Examples include "What was good enough for Daddy is good enough for me" and "My model is nicer than your model". One "game," in particular, named "The delusions we suffer," described by Dunnette as "probably the most dangerous game of all" (p. 346), always caught my attention for its pertinence to the goings-on I was noticing in the literature. According to Dunnette's description, this "game" consists of "maintaining delusional systems to support our claims that the things we are doing *really* constitute good science". Dunnette went on to chastise "a pattern of self-deceit which becomes more ingrained and less tractable with each new delusion" (p. 346).

One of the numerous variants of this pursuit of ways to keep our delusions alive is the "[rationalization of] certain practices on the grounds that they are intrinsically good for humanity and that they need not, therefore, meet the usual standards demanded by scientific verification" (p. 346). Another variant is the "search for General Laws," in which "individual differences [are treated] as merely bothersome variation – to be reduced by adequate controls or treated as error variance" (p. 347). On this issue, Dunnette commented that "we cannot expect a science of human behavior to advance far until the moderating effects of individual variation on the functional relationships being studied are taken fully into account. People do, after all, differ greatly from one another" (p. 347). Both variants of "delusions we suffer" can be found in research on the relationship between physical activity and mental health (Biddle & Ekkekakis, 2005; Ekkekakis & Backhouse, 2009).

1. Delusion A: exercise effects "need not meet the standards of scientific verification"

Exercise is a quintessential example of a practice that is seen, especially by its scientists, practitioners, and other enthusiasts, as "intrinsically good for humanity," to use Dunnette's words. That exercise makes people "feel better" is such a positive and uplifting message that most people who are passionate about exercise are probably eager to shout it from the rooftops. The problem is that when well-intentioned researchers are caught up in the spell-binding allure of their own convictions, they tend to adopt a set to prove rather than disprove their study hypotheses. Arguably,

researchers in the exercise sciences have a spotty track record of sacrificing scientific impartiality at the altar of promoting exercise as a healthful panacea. As Salmon (2001) pointed out, claims for the mental health benefits of exercise "have tended to anticipate rather than reflect the accumulation of strong evidence" (p. 36). As an example, at a time when the published studies were few and methodologically weak, Morgan (1981) confidently proclaimed that "the 'feeling better' sensation that accompanies regular physical activity is so obvious that it is one of the few universally accepted benefits of exercise" (p. 306).

The accumulation of evidence over the past three decades has enabled an expert panel to arrive at the conclusion that the extant evidence, at least in the case of depression, suffices to satisfy the traditional criteria for the establishment of causation. Specifically, Biddle, Fox, Boutcher, and Faulkner (2000) asserted that "overall, the evidence is strong enough for us to conclude that there is support for a causal link between physical activity and reduced clinically defined depression. This is the first time such a statement has been made" (p. 155). However, not everyone agrees. In a meta-regression analysis on the effects of exercise on depression, which some have since criticized as "a bit harsh" (Brosse, Sheets, Lett, & Blumenthal, 2002; p. 754) and others as polemical (Landers & Arent, 2007), Lawlor and Hopker (2001) coded for study quality based on whether they found indications of (a) adequate concealment of group allocation, (b) intention-to-treat analyses, and (c) blinding of outcome assessors to study hypotheses. Adequate concealment, in particular, was defined as "central randomization at a site remote from the study; computerized allocation in which records are in a locked, unreadable file that can be accessed only after entering patient details; the drawing of sealed and opaque sequentially numbered envelopes" (p. 2). Although these may seem as inconsequential details, they give a glimpse of the methodological rigor that critical reviewers have come to expect of all randomized clinical trials. Of the 14 studies examined by Lawlor and Hopker, randomization was concealed in only 3, intention-to-treat analyses were undertaken in only 2, and assessment of outcome was blinded in only 1. None of the 14 studies satisfied all three of these quality criteria, leading Lawlor and Hopker (2001) to conclude that "most studies were of poor quality" (p. 3).

Concealment, intention to treat analyses, and blinding certainly do not exhaust the list of crucial methodological elements that could influence the outcome of a study (perhaps significantly more than the opacity of an envelope). As a case in point, consider a study that is widely regarded as one of the most complete empirical examinations of the effects of exercise on depression to date. In the original report, Blumenthal et al. (1999) described participant

recruitment in these words: “Subjects were recruited through flyers, media advertisements, and letters sent to local physicians and mental health facilities” (p. 2350). Yet, in the discussion of the 10-month follow-up, *Babyak et al. (2000)* added some important details: “The sample consisted of patient – volunteers who responded to advertisements seeking participants for a study of exercise therapy for depression. We presume that these participants believed exercise to be a credible treatment modality for depression and were favorably inclined toward participation” (p. 637). Certainly, it does not take a reviewer with an anti-exercise bias to detect the potential for distortion created when someone volunteers for “a study of exercise therapy for depression”. Portraying exercise as “therapy for depression” in advertising materials is not much different from the method used by *Desharnais, Jobin, Côté, Lévesque, and Godin (1993)* to deliberately create an expectation of benefit; they told their participants that “the training program was designed to improve both aerobic capacity and psychological well-being” (p. 151). It is reasonable to expect that a volunteer for “a study of exercise therapy for depression,” if he or she is assigned to the exercise group, might tend to exhibit a favorable attitude toward exercise, particularly considering that most symptoms of depression are subjective and the methods of assessment (questionnaire and clinical interview) are based on self-report (recall, for example, that, in a recent study, remission with exercise treatments reached 40–45%, while that associated with placebo was 31%; *Blumenthal et al., 2007*). Conversely, if a volunteer who signs up for “a study of exercise therapy for depression” is assigned to a non-exercise group (e.g., pharmacotherapy or wait-list), it is reasonable to expect a certain degree of discontent. *Babyak et al. (2000)* commented that “it was apparent that there may have been some ‘antimedication’ sentiment among some study participants” (p. 636). Some of those who were assigned to the combined exercise-plus-sertraline group expressed “disappointment” and “mentioned spontaneously that the medication seemed to interfere with the beneficial effects of the exercise program” (p. 636). Almost half (48%) of the participants assigned to the sertraline group initiated an exercise program on their own after the 16-week treatment, compared to the 6% of participants originally assigned to the exercise group who chose to initiate pharmacotherapy after the end of the treatment period.

Methodological weaknesses and interpretational pitfalls such as these and numerous others underscore the great challenges involved in establishing causation in human studies investigating the effects of exercise on mental health. It was recently argued that, for a treatment to be supported by what has been described as “Level 1, Grade A” evidence (*Guyatt, Cook, Sackett, Eckman, & Pauker, 1998*), the following criteria must be met: (a) the evidence must come from randomized, controlled trials, (b) there must be a high benefit-to-risk ratio, (c) the results must be clear-cut, and (d) the evidence must come from a large sample (*Wipfli, Rethorst, & Landers, 2008*). It was also argued that, since “large-scale randomized trials are often extremely costly and time consuming,” one can take several “Level 2” studies and combine them through a meta-analysis, to arrive at “Level 1, Grade A” evidence (*Wipfli et al., 2008*; p. 394). In actuality, *Guyatt et al. (1998)* proposed that what distinguishes “Level 1, Grade A” from “Level 2, Grade A” evidence is not the size of the sample but rather that for “Level 1,” the effect is unambiguous and the benefits clearly outweigh the risks, whereas for “Level 2,” the effect is equivocal and there is uncertainty as to whether the benefits outweigh the risks. Nevertheless, *Wipfli et al. (2008)*, finding an average effect size of -0.48 from “randomized controlled trials” examining the effects of exercise on anxiety (parenthetically, mixing both acute exercise bouts and chronic exercise programs in a classic “apples and oranges” scenario; see *Eysenck, 1994, 1995*) concluded that “these results provide Level 1, Grade A evidence for using exercise in the

treatment of anxiety” (p. 392). This conclusion was reached despite acknowledging that, in nearly all (46 of 49) of the studies reviewed, the participants were not clinically anxious and, therefore, anxiety was not really examined as *treatment* for anxiety. One major point that apparently escaped the attention of *Wipfli et al. (2008)* is that, according to *Guyatt et al. (1998)*, for Grade A (and even Grade B) evidence, the methods must be “strong” (p. 442S). However, *Wipfli et al. (2008)* did not code for methodological quality and/or threats to internal validity, relying instead on the debatable argument that the studies that were included in the meta-analysis were “high-quality studies” (p. 402) by virtue of being “randomized controlled trials”. Yet, in only 12 of these 49 “high-quality studies” was there for sufficient information about the amount of exercise the participants performed (i.e., intensity, duration)! As discussed in the previous paragraph, the presence of a control group and the random allocation of participants to treatment groups clearly do not suffice to characterize a study as “high-quality,” since numerous other factors can threaten internal validity and can undermine the establishment of a causal relationship between the independent and the dependent variable.

Guidelines for evaluating the quality of randomized controlled trials included in meta-analyses and systematic reviews have been published (e.g., *Jüni, Altman, & Egger, 2001*) and most major journals in the field of health care have now adopted reporting guidelines (e.g., *Boutron, Moher, Altman, Schulz, & Ravaut, 2008*). Establishing causation in exercise studies (let alone developing an adequate evidence base to justify public proclamations about using exercise as treatment for serious mental health problems) is an extremely difficult task. It is made even more challenging by the fact that there can be no placebo exercise control; participants in the exercise group will always be aware that they were assigned to the exercise group. This inescapable situation should make researchers even more cautious and conservative in their interpretations and conclusions. Criticism is not to be dismissed or scorned but carefully contemplated. Notice, for example, that after *Lawlor and Hopker (2001)* specified that they evaluated adequate concealment by whether or not “sealed and opaque sequentially numbered envelopes” (p. 2) were drawn, *Dunn, Trivedi, Kampert, Clark, and Chambliss (2005)* complied, stating that “randomization was implemented with sequentially numbered, opaque, sealed envelopes” (p. 2). More importantly, after *Lawlor and Hopker (2001)* speculated that “socializing” could mediate the effects of exercise, since “none of the participants in the studies [they] reviewed exercised alone” (p. 5), *Dunn et al. (2005)* addressed the criticism by having their participants exercise “individually in rooms by themselves” (p. 2). Arguably, this is how progress is made: through the cycle of criticism and methodological improvement.

For many researchers, whether exercise is causally related to positive mental health changes remains an open question and rightly so. Effect sizes of 1.1 (*Lawlor & Hopker, 2001*) or 1.4 (*Stathopoulou, Powers, Berry, Smits, & Otto, 2006*) based on randomized clinical trials are impressive but, by themselves, do not suffice. The proverbial “devil” usually resides in the methodological details, many of which (if disclosed) are not quite as impressive. Given the inability to devise an exercise placebo control, perhaps the most convincing evidence for causation that has emerged so far has come from basic animal research (e.g., *Bjørnebekk, Mathé, & Brené, 2005*; *Duman, Schlesinger, Russell, & Duman, 2008*; *Greenwood, Strong, Brooks, & Fleshner, 2008*), a field, of course, with its own unique interpretational challenges.

2. Delusion B: individual differences in response to exercise are “merely bothersome variation”

True to our statistical training in the grand tradition of the general linear model, most of us rarely examine individual patterns

and trends in our data, focusing all our attention instead on the venerable *p* value. At best, we try to get rid of the annoying individual differences by including covariates in our general linear models. References to the role of personality or temperament have become exceedingly rare. A recent meta-analysis on the role of personality in physical activity uncovered only 33 studies published over a period of 27 years (Rhodes & Smith, 2006). Naturally, as more generations of new scientists are being educated in a research culture devoid of reminders about the prevalence and importance of individual differences, a tacit assumption begins to develop that such differences either do not exist or do not matter.

When we first published a paper showing pronounced inter-individual variability in affective responses to a bout of aerobic exercise (Van Landuyt, Ekkekakis, Hall, & Petruzzello, 2000), the feedback we received ranged from incredulous to angry (because the finding implied that, contrary to what conventional wisdom in our field would have predicted, some participants did not “feel better” during moderate-intensity exercise). After I presented other variability data at a conference (e.g., see Ekkekakis, Hall, & Petruzzello, 2005a), one senior colleague publicly admonished me by saying that examining individual differences “is unscientific” and that we should strive to find a global dose–response relationship “the same way they do it in biology”. Although I had made it clear that we had carefully controlled for differences in aerobic fitness either by fixing intensity at the same percentage of maximal aerobic capacity or the same percentage of oxygen uptake associated with the ventilatory threshold, another colleague, snickering contemptuously (I still have the tape!), asked me whether we had considered if “all this variability” was “just due to differences in fitness” (which was followed by audible laughter from a portion of the audience). Still to this day, I get comments from reviewers saying that there is no evidence that the affective response to the same exercise intensity varies significantly between individuals.

Recently, we proposed the constructs of preference for exercise intensity and tolerance of exercise intensity. We conceptualized these as being manifestations of individual differences in somatosensory modulation and as having a genetic basis, writing, for example (see Ekkekakis, Hall, & Petruzzello, 2005b for the references omitted from this excerpt):

There is evidence that, much like heritable variation in extraversion and sensation-seeking and pain sensitivity and tolerance, the variation in the preferred intensity of exercise might also be partly heritable. This is supported by human twin studies, human DNA studies, animal DNA studies, and animal artificial selection studies. Although the bulk of genotypic differences are linked to peripheral physiological traits, such as muscle glucose uptake, studies have also highlighted the important role of brain regulatory mechanisms (pp. 355–356).

Predictably, a reviewer told us “I am not sure that these are predispositions (as argued in the paper) as much as they are experiential results”. This led to the inclusion of the following caveat, which made up in political correctness what it lacked in insight:

We emphasize that we do not consider these traits to be the sole determinants of intensity selection or tolerance. Other factors including physical (e.g., fitness, age, health status), experiential (e.g., learned coping skills, exercise history), and situational (e.g., self-efficacy, social physique anxiety) are also likely to be important (p. 354).

The 8-item scale we developed to assess the construct of preference for exercise intensity was since found to account for 17–18% of the variance in self-selected intensity (percentage of the oxygen uptake associated with the ventilatory threshold), beyond the 16–20% accounted for by the combination of age, body mass index,

and maximal aerobic capacity (Ekkekakis, Lind, & Joens-Matre, 2006). The 8-item scale developed to assess tolerance of exercise intensity accounted for 14% of the variance in the amount of time young men and women persisted during a graded treadmill test to volitional termination after they exceeded their ventilatory threshold, beyond the 13% accounted for by age, body mass index, frequency of habitual physical activity, and habitual duration of physical activity per session. In a different sample of sedentary middle-aged women, the tolerance scale accounted for 20% of the variance, beyond the 42% contributed by the combination of age, body mass index, and maximal aerobic capacity (Ekkekakis, Lind, Hall, & Petruzzello, 2007).

Although rarely the focus of high-profile research, evidence for large variability in various responses to exercise behavior across the animal kingdom abounds in the literature, if one cares to look. Importantly, this variability cannot be accounted for by differences in morphological and physiological characteristics.

With regard to preference for physical activity, Premack and Schaeffer (1963) found that, among six female albino rats, one averaged 46 wheel revolutions per hour across multiple observation periods, whereas another averaged 515 revolutions per hour (an 11-fold difference). On one of the observation days, the difference was as large as 15 versus 556 revolutions per hour (a 37-fold difference). In their observations of the physical activity of 18 adult female rhesus monkeys using accelerometers, Sullivan, Koegler, and Cameron (2006) found an 8-fold difference in daily activity counts between the most and the least active animal. Importantly, activity patterns were consistent within each animal over time (3-month test-retest, $r = .79$). In sedentary middle-aged women, we have found that the range of self-selected intensity of walking extended from 60 to 160% of the level of oxygen uptake associated with the ventilatory threshold (Lind, Joens-Matre, & Ekkekakis, 2005).

With regard to exercise tolerance, the differences seem even larger. A study of endurance performance in lizards found a 47-fold range, from 36 to 1677 s, a range that could not be explained by differences in physiological capacity (Le Galliard, Clobert, & Ferrière, 2004). In rats, Bedford, Tipton, Wilson, Opplinger, and Gisolfi (1979) noted that “approximately 10% of mature rats will not run, or continue to run, when placed on the treadmill” (p. 1278) but an examination of a range of physiological parameters showed that “no statistically significant differences between animals willing and unwilling to run” (p. 1279). In mice, Dohm, Hayes, and Garland (1996) found that although all animals swam when placed in the water, there was a 386-fold range in performance among females (from 0.72 to 278.2 min) and a 395-fold range among males (from 0.42 to 165.9 min). Importantly, swimming endurance was reproducible on two occasions ($r = .64$), did not change significantly from the first to the second trial, and was unrelated to body mass and maximal oxygen uptake. In humans, it has been reported that, contrary to what is commonly assumed in exercise physiology, only as few as 17% of participants reach a “plateau” in oxygen consumption (i.e., “max-out” physiologically) during incremental exercise tests to volitional termination (Day, Rossiter, Coats, Skasick, & Whipp, 2003). Accordingly, Gulati et al. (2005) reported that, in their cohort of 5721 asymptomatic women, the range in the percentage of age-predicted exercise capacity actually achieved during a “symptom-limited” treadmill test extended from 20 to 150%.

Given the findings of large inter-individual differences (beyond what can be accounted for by differences in morphological characteristics and living conditions) and intra-individual consistency across a range of species, it should not be surprising that genetic factors have attracted research attention. In their extensive study across 13 strains of mice, Lightfoot, Turner, Daves, Vordermark,

and Kleeberger (2004) found that heritability estimates for wheel-running distance were 31–48% for males and 12–22% for females; for duration, 44–61% for males and 12–21% for females; and for velocity, 49–66% for males and 44–61% for females. Given these figures, it is also not surprising that artificial selection experiments in both mice (Swallow, Carter, & Garland, 1998) and rats (Morishima-Yamato et al., 2005), consisting of pairing male and female individuals with a high wheel-running propensity, have been successful in producing offspring exhibiting much higher levels of daily activity than randomly paired controls.

Importantly, the offspring in these artificial selection experiments achieve the higher levels of activity primarily by running at higher average speeds (i.e., at a higher intensity, consistent with the finding of higher heritability estimates for running velocity than distance and duration by Lightfoot et al., 2004). Commenting on these findings, Rezende, Chappell, Gomes, Malisch, and Garland (2005) noted that the mice from the selection lines never actually reach their maximal aerobic capacity while running voluntarily and their maximal aerobic or endurance capacity does not seem to be associated with or to limit their voluntary running. Instead, what drives the selected mice to run more, mainly by running faster, according to the authors, is a “genetically higher motivation for wheel-running” (p. 2447). It is interesting to point out that, consistent with the results of animal studies, heritability indices in human studies have also tended to be somewhat higher for the more vigorous forms of physical activity, such as intense exercise or sports (Beunen & Thomis, 1999; Lauderdale et al., 1997; Maia, Thomis, & Beunen, 2002). This suggests that genetic factors may become more relevant when the activity stimulus is more (physically or psychologically) demanding.

The next obvious challenge for researchers is to identify the genes responsible for these inter-individual differences. Genome-wide scans for quantitative trait loci (QTL), which had been initiated in the last decade (e.g., Gershenfeld et al., 1997; Mayeda & Hofstetter, 1999), are now continuing for both running behavior (Lightfoot, Turner, Pomp, Kleeberger, & Leamy, 2008) and endurance performance (Lightfoot et al., 2007). Ideally, however, given the susceptibility of QTL analyses to Type I and Type II errors, a more focused, hypothesis-driven approach would be preferable. From a psychological perspective, the interest is on genes that may account for differences in spontaneous behavior, motivation, or tolerance. Unfortunately, on this issue, progress has been slower than one would have hoped. Efforts to explore the genetic bases of variation in exercise performance have focused overwhelmingly on such factors as energy supply and oxygen transport. In an annually updated series of articles examining genes associated with exercise performance and fitness, the first reference to “behavioral” factors was made only in the fifth installment (Wolfarth et al., 2005). According to this review, only 5 of 230 relevant papers published in 2004 (approximately 2%) referred to “behavior”.

Perhaps the most interesting approach so far from a psychological standpoint has been driven by the hypothesis that a tendency for higher levels of physical activity must be associated with neurotransmitter systems, such as dopamine, which are involved in appetitive behaviors, approach tendencies, and natural reward. Using mice from their artificial selection experiment, Rhodes, Garland, and Gammie (2003) identified the nucleus accumbens (considered a part of the dopamine reward system) as an area on which artificial selection had acted to produce the observed differences in wheel-running behavior between experimental and control animals. Rhodes, Gammie, and Garland (2005) concluded that “the major changes in High-Runner lines appear to have taken place in the brain rather than in capacities for exercise” (p. 438). Following a very similar rationale but using two large human cohorts, Simonen et al. (2003) found a link (particularly consistent

among White women) between a polymorphism in a gene that encodes for the D2 type of dopamine receptor (DRD2) and physical activity participation over the previous year.

The research linking genes to behavioral characteristics is recent, having started only in 1996 (with the publication of two studies in the same issue of *Nature Genetics*, linking a polymorphism of the dopamine D4 receptor, DRD4, to novelty-seeking behavior). Since then, this research has exploded but the picture remains murky as substantial challenges are rising to the surface (Kagan, 2007). At this point, the evidence that is of interest to researchers focusing on physical activity behavior and the effects of physical activity on mental health suggests a weak but reliable association between a DRD4 polymorphism and the approach-related traits of novelty seeking and impulsivity (Munafò, Yalcin, Willis-Owen, & Flint, 2008), as well as a promising connection between a polymorphism of the gene that encodes for the serotonin transporter (5-HTT) with emotion regulation (Canli & Lesch, 2007; Hariri & Holmes, 2006).

Of interest, particularly to researchers focusing on individual differences in tolerance to high-intensity exercise, is also a polymorphism of the catechol-O-methyltransferase (COMT) gene, which codes the substitution of valine by methionine at codon 158 (*val¹⁵⁸met*). COMT is one of the enzymes involved in catecholamine degradation and, therefore, has a key modulatory role in dopaminergic neurotransmission. Individuals with the *met/met* genotype show the lowest COMT activity and, therefore, exhibit chronic hyperactivity of the dopaminergic system. This, then, is linked to a reduction in enkephalin peptides and a compensatory increase in mu-opioid receptor concentrations. Conversely, individuals with the *val/val* genotype have the highest COMT activity and, therefore, show reduced dopaminergic neurotransmission. Consequently, such individuals also show increased enkephalin and reduced mu-opioid receptor concentrations. A fascinating study by Zubieta et al. (2003) found that *met/met* homozygotes showed diminished opioid system responses to painful stimulation, higher sensory and affective pain ratings, and a more negative affective state. A recent genetic imaging study by Smolka et al. (2007) showed an additive effect of the COMT and 5-HTT gene polymorphisms, accounting for 40% of the inter-individual variance in the response of limbic structures (such as the amygdala and hippocampus) to unpleasant (in this case, visual) stimuli, assessed by functional magnetic resonance imaging. It is probably reasonable to predict that studies investigating the relationship of the DRD4, COMT, and 5-HTT gene polymorphisms with physical activity behavior will emerge in the future, although how long the waiting period will be seems unpredictable.

Thanks mainly to Claude Bouchard's authoritative influence (e.g., Bouchard & Rankinen, 2001), exercise scientists are slowly becoming sensitized to the importance of individual differences in exercise adaptability, as well as to the importance of genetic factors. There are now studies specifically focused on individual differences in responses to exercise training (e.g., Hautala et al., 2006), which would have been hard to imagine in the past. An intriguing recent study examined individual differences in exercise-induced weight loss, a phenomenon of great public interest but one that had received very little research attention until now (King, Hopkins, Caudwell, Stubbs, & Blundell, 2008). Perhaps, with any luck, one of these days, the reviewers of my manuscripts will also come to terms with the fact that not everyone experiences the same increase in pleasure in response to a bout of exercise. Perhaps some day, in the not-so-distant future, we may even begin to explore whether such inter-individual differences in affective responses have something to do with exercise behavior and adherence (Backhouse, Ekkekakis, Biddle, Foksett, & Williams, 2007; Williams et al., 2008).

3. Eco's deafening echoes

In this issue of the journal, Eco de Geus and his collaborator Marleen de Moor present an exceptionally provocative article. Its bold purpose, based on a series of studies on twins, is nothing less than to question two of our most “ingrained” and least “tractable” delusions: (a) that the positive *association* between exercise and various components of mental health reflects *causal* effects of exercise and (b) that exercise exerts beneficial effects of more-or-less equal magnitude on *all* participants. The authors' argument is twofold. First, they claim that, at least in part, some common genetic factors may influence both the tendency for physical activity behavior and the propensity of some individuals for good physical and mental health. Second, they argue that there are large individual differences not only in the magnitude of the exercise-associated changes but also in the direction, with some responding favorably, some experiencing no response, and others responding unfavorably.

Not only do de Geus and de Moor present intriguing evidence in support of their claims but they also put forth a conceptual model designed to answer what they rightly call the two “most vexing questions in the field of exercise intervention: why exercisers exercise, and why non-exercisers do not”. Their idea is that well-being results mainly from the “feel-better” effects of acute exercise bouts and the boost in self-esteem experienced by those who adapt to exercise well and derive visible performance benefits. Individuals who do “feel better” in response to exercise and do see improvements in exercise capacity are more likely to continue their exercise participation. On the other hand, those for whom the exercise stimulus is aversive and/or fail to experience visible performance gains are more likely to drop out. Where genetic influences come in is by predisposing some individuals to (a) experience improved affect in response to exercise bouts and (b) experience greater gains in exercise capacity with training.

Although they make a strong case for genetic influences on exercise behavior, de Geus and de Moor vehemently reject narrow notions of genetic determinism. In their implications for exercise interventions, they emphasize that what their model means is that “we should not close our eyes to human genetic variation” and should instead acknowledge that “it may be harder to engage some people in exercise than others”. For those who are harder to engage, an individualized approach to exercise prescription might be indicated, one that “emphasizes the appetitive aspects” and “reduces the aversive aspects”.

It should be evident by now that I consider the article by de Geus and de Moor to be a work of exceptional importance and great potential impact for the field of physical activity and mental health. Others may agree or disagree. For example, in a recent debate in the *Journal of Applied Physiology*, Roth (2008), aligned with de Geus and de Moor, suggests that the current approach to exercise prescription may not be optimal for some individuals and, consequently, he envisions “the future application of genomic information to improve and individualize exercise prescription” (p. 1244). However, several of the respondents disagree on various grounds. Those adhering to the “orthodox” view counter, for example, that individual differences, such as those that have been presented by Bouchard, perhaps represent “[nothing] more than random noise” or “random measurement variance” and conclude that it would be “important to identify non-genomic explanations for the negative responses before blaming genes” (Wagner, 2008; p. 1246). Certainly, old paradigms are not replaced overnight. So, whatever anyone may believe, at the very least, I hope that everyone will take a minute to study the arguments and contemplate the validity of our most “ingrained” and least “tractable” delusions.

References

- Babyak, M., Blumenthal, J. A., Herman, S., Khatri, P., Doraiswamy, M., Moore, K., et al. (2000). Exercise treatment for major depression: maintenance of therapeutic benefit at 10 months. *Psychosomatic Medicine*, 62, 633–638.
- Backhouse, S. H., Ekkekakis, P., Biddle, S. J. H., Foskett, A., & Williams, C. (2007). Exercise makes people feel better but people are inactive: paradox or artifact? *Journal of Sport and Exercise Psychology*, 29, 498–517.
- Bedford, T. G., Tipton, C. M., Wilson, N. C., Opplinger, R. A., & Gisolfi, C. V. (1979). Maximum oxygen consumption of rats and its changes with various experimental procedures. *Journal of Applied Physiology*, 47, 1278–1283.
- Beunen, G., & Thomis, M. (1999). Genetic determinants of sports participation and daily physical activity. *International Journal of Obesity*, 23(Suppl. 3), S55–S63.
- Biddle, S. J. H., & Ekkekakis, P. (2005). Physically active lifestyles and well-being. In F. A. Huppert, B. Keverne, & N. Baylis (Eds.), *The science of well-being* (pp. 140–168). Oxford, United Kingdom: Oxford University Press.
- Biddle, S. J. H., Fox, K. R., Boutcher, S. H., & Faulkner, G. E. (2000). The way forward for physical activity and the promotion of psychological well-being. In S. J. H. Biddle, K. R. Fox, & S. H. Boutcher (Eds.), *Physical activity and psychological well-being* (pp. 154–168). London: Routledge.
- Bjørnebekk, A., Mathé, A. A., & Brené, S. (2005). The antidepressant effect of running is associated with increased hippocampal cell proliferation. *International Journal of Neuropsychopharmacology*, 8, 357–368.
- Blumenthal, J. A., Babyak, M. A., Doraiswamy, P. M., Watkins, L., Hoffman, B. M., Barbour, K. A., et al. (2007). Exercise and pharmacotherapy in the treatment of major depressive disorder. *Psychosomatic Medicine*, 69, 587–596.
- Blumenthal, J. A., Babyak, M. A., Moore, K. A., Craighead, W. E., Herman, S., Khatri, P., et al. (1999). Effects of exercise training on older patients with major depression. *Archives of Internal Medicine*, 159, 2349–2356.
- Bouchard, C., & Rankinen, T. (2001). Individual differences in response to regular physical activity. *Medicine and Science in Sports and Exercise*, 33, S446–S451.
- Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., & Ravaut, P. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine*, 148, 295–309.
- Brosse, A. L., Sheets, E. S., Lett, H. S., & Blumenthal, J. A. (2002). Exercise and the treatment of clinical depression in adults. Recent findings and future directions. *Sports Medicine*, 32, 741–760.
- Canli, T., & Lesch, K.-P. (2007). Long story short: the serotonin transporter in emotion regulation and social cognition. *Nature Neuroscience*, 10, 1103–1109.
- Day, J. R., Rossiter, H. B., Coats, E. M., Skasick, A., & Whipp, B. J. (2003). The maximally attainable VO₂ during exercise in humans: the peak vs. maximum issue. *Journal of Applied Physiology*, 95, 1901–1907.
- Desharnais, R., Jobin, J., Côté, C., Lévesque, L., & Godin, G. (1993). Aerobic exercise and the placebo effect: a controlled study. *Psychosomatic Medicine*, 55, 149–154.
- Dohm, M. R., Hayes, J. P., & Garland, T., Jr. (1996). Quantitative genetics of sprint running speed and swimming endurance in laboratory house mice (*Mus domesticus*). *Evolution*, 50, 1688–1701.
- Duman, C. H., Schlesinger, L., Russell, D. S., & Duman, R. S. (2008). Voluntary exercise produces antidepressant and anxiolytic behavioral effects in mice. *Brain Research*, 1199, 148–158.
- Dunn, A. L., Trivedi, M. H., Kampert, J. B., Clark, C. G., & Chambliss, H. O. (2005). Exercise treatment for depression: efficacy and dose-response. *American Journal of Preventive Medicine*, 28, 1–8.
- Dunnette, M. D. (1966). Fads, fashions, and folderol in psychology. *American Psychologist*, 21, 343–352.
- Ekkekakis, P., & Backhouse, S. H. (2009). Exercise and psychological well-being. In R. Maughan (Ed.), *Olympic textbook of science in sport* (pp. 249–269). Chichester, United Kingdom: Wiley-Blackwell.
- Ekkekakis, P., Hall, E. E., & Petruzzello, S. J. (2005a). Variation and homogeneity in affective responses to physical activity of varying intensities: an alternative perspective on dose-response based on evolutionary considerations. *Journal of Sports Sciences*, 23, 477–500.
- Ekkekakis, P., Hall, E. E., & Petruzzello, S. J. (2005b). Some like it vigorous: individual differences in the preference for and tolerance of exercise intensity. *Journal of Sport and Exercise Psychology*, 27, 350–374.
- Ekkekakis, P., Lind, E., Hall, E. E., & Petruzzello, S. J. (2007). Can self-reported tolerance of exercise intensity play a role in exercise testing? *Medicine and Science in Sports and Exercise*, 39, 1193–1199.
- Ekkekakis, P., Lind, E., & Joens-Matre, R. R. (2006). Can self-reported preference for exercise intensity predict physiologically defined self-selected exercise intensity? *Research Quarterly for Exercise and Sport*, 77, 81–90.
- Eysenck, H. J. (1994). Systematic reviews: meta-analysis and its problems. *British Medical Journal*, 309, 789–792.
- Eysenck, H. J. (1995). Meta-analysis or best-evidence synthesis? *Journal of Evaluation in Clinical Practice*, 1, 29–36.
- Gershenfeld, H. K., Neumann, P. E., Mathis, C., Crawley, J. N., Li, X., & Paul, S. M. (1997). Mapping quantitative trait loci for open-field behavior in mice. *Behavior Genetics*, 27, 201–210.
- Greenwood, B. N., Strong, P. V., Brooks, L., & Fleshner, M. (2008). Anxiety-like behaviors produced by acute fluoxetine administration in male Fischer 344 rats are prevented by prior exercise. *Psychopharmacology*, 199, 209–222.
- Gulati, M., Black, H. R., Shaw, L. J., Arnsdorf, M. F., Bairey Merz, C. N., Lauer, M. S., et al. (2005). The prognostic value of a nomogram for exercise capacity in women. *New England Journal of Medicine*, 353, 468–475.

- Guyatt, G. H., Cook, D. J., Sackett, D. L., Eckman, M., & Pauker, S. (1988). Grades of recommendation for antithrombotic agents. *Chest*, *114*, 4415–4445.
- Hariri, A. R., & Holmes, A. (2006). Genetics of emotional regulation: the role of the serotonin transporter in neural function. *Trends in Cognitive Sciences*, *10*, 182–191.
- Hautala, A. J., Kiviniemi, A. M., Mäkikallio, T. H., Kinnunen, H., Nissilä, S., Huikuri, H. V., et al. (2006). Individual differences in the responses to endurance and resistance training. *European Journal of Applied Physiology*, *96*, 535–542.
- Jüni, P., Altman, D. G., & Egger, M. (2001). Assessing the quality of controlled clinical trials. *British Medical Journal*, *323*, 42–46.
- Kagan, J. (2007). A trio of concerns. *Perspectives on Psychological Science*, *2*, 361–376.
- King, N. A., Hopkins, M., Caudwell, P., Stubbs, R. J., & Blundell, J. E. (2008). Individual variability following 12 weeks of supervised exercise: identification and characterization of compensation for exercise-induced weight loss. *International Journal of Obesity*, *32*, 177–184.
- Landers, D. M., & Arent, S. M. (2007). Physical activity and mental health. In G. Tenenbaum, & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed.). (pp. 469–491) New York: Wiley & Sons.
- Lauderdale, D. S., Fabsitz, R., Meyer, J. M., Sholinsky, P., Ramakrishnan, V., & Goldberg, J. (1997). Familial determinants of moderate and intense physical activity: a twin study. *Medicine and Science in Sports and Exercise*, *29*, 1062–1068.
- Lawlor, D. A., & Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised control trials. *British Medical Journal*, *322*, 1–8.
- Le Galliard, J.-F., Clobert, J., & Ferrière, R. (2004). Physical performance and Darwinian fitness in lizards. *Nature*, *432*, 502–505.
- Lightfoot, J. T., Turner, M. J., Daves, M., Vordermark, A., & Kleeberger, S. R. (2004). Genetic influence on daily wheel running activity level. *Physiological Genomics*, *19*, 270–276.
- Lightfoot, J. T., Turner, M. J., Kleinfehn Knab, A., Jedlicka, A. E., Oshimura, T., Marzec, J., et al. (2007). Quantitative trait loci associated with maximal exercise endurance in mice. *Journal of Applied Physiology*, *103*, 105–110.
- Lightfoot, J. T., Turner, M. J., Pomp, D., Kleeberger, S. R., & Leamy, L. J. (2008). Quantitative trait loci for physical activity traits in mice. *Physiological Genomics*, *32*, 401–408.
- Lind, E., Joens-Matre, R. R., & Ekkekakis, P. (2005). What intensity of physical activity do formerly sedentary middle-aged women select? Evidence of a coherent pattern from physiological, perceptual, and affective markers. *Preventive Medicine*, *40*, 407–419.
- Maia, J. A. R., Thomis, M., & Beunen, G. (2002). Genetic factors in physical activity levels: a twin study. *American Journal of Preventive Medicine*, *23*(Suppl. 2), 87–91.
- Mayeda, A. R., & Hofstetter, J. R. (1999). A QTL for the genetic variance in free-running period and level of locomotor activity between inbred strains of mice. *Behavior Genetics*, *29*, 171–176.
- Morgan, W. P. (1981). Psychological benefits of physical activity. In F. J. Nagle, & H. J. Montoye (Eds.), *Exercise in health and disease* (pp. 299–314). Springfield, IL: Charles C. Thomas.
- Morishima-Yamato, M., Hisaoka, F., Shinomiya, S., Harada, N., Matoba, H., Takahashi, A., et al. (2005). Cloning and establishment of a line of rats for high levels of voluntary wheel running. *Life Sciences*, *77*, 551–561.
- Munafò, M. R., Yalcin, B., Willis-Owen, S. A., & Flint, J. (2008). Association of the dopamine D4 receptor (DRD4) gene and approach-related personality traits: meta-analysis and new data. *Biological Psychiatry*, *63*, 197–206.
- Premack, D., & Schaeffer, R. W. (1963). Some parameters affecting the distributional properties of operant-level running in rats. *Journal of the Experimental Analysis of Behavior*, *6*, 473–475.
- Rezende, E. L., Chappell, M. A., Gomes, F. R., Malisch, J. L., & Garland, T. (2005). Maximal metabolic rates during voluntary exercise, forced exercise, and cold exposure in house mice selectively bred for high wheel-running. *Journal of Experimental Biology*, *208*, 2447–2458.
- Rhodes, J. S., Gammie, S. C., & Garland, T. (2005). Neurobiology of mice selected for high voluntary wheel-running activity. *Integrative Comparative Biology*, *45*, 438–455.
- Rhodes, J. S., Garland, T., & Gammie, S. C. (2003). Patterns of brain activity associated with variation in voluntary wheel-running behavior. *Behavioral Neuroscience*, *117*, 1243–1256.
- Rhodes, R. E., & Smith, N. E. I. (2006). Personality correlates of physical activity: a review and meta-analysis. *British Journal of Sports Medicine*, *40*, 958–965.
- Roth, S. M. (2008). Viewpoint: perspective on the future use of genomics in exercise prescription. *Journal of Applied Physiology*, *104*, 1243–1245.
- Salmon, P. (2001). Effects of physical exercise on anxiety, depression, and sensitivity to stress: a unifying theory. *Clinical Psychology Review*, *21*, 33–61.
- Simonen, R. L., Rankinen, T., Pérusse, L., Leon, A. S., Skinner, J. S., Wilmore, J. H., et al. (2003). A dopamine D2 receptor gene polymorphism and physical activity in two family studies. *Physiology and Behavior*, *78*, 751–757.
- Smolka, M. N., Bühler, M., Schumann, G., Klein, S., Hu, X.-Z., Moayer, M., et al. (2007). Gene-gene effects on central processing of aversive stimuli. *Molecular Psychiatry*, *12*, 307–317.
- Stathopoulou, G., Powers, M. B., Berry, A. C., Smits, A. J., & Otto, M. W. (2006). Exercise interventions for mental health: a quantitative and qualitative review. *Clinical Psychology: Science and Practice*, *13*, 179–193.
- Sullivan, E. L., Koegler, F. H., & Cameron, J. L. (2006). Individual differences in physical activity are closely associated with changes in body weight in adult female rhesus monkeys (*Macaca mulatta*). *American Journal of Physiology*, *291*, R633–R642.
- Swallow, J. G., Carter, P. A., & Garland, T. (1998). Artificial selection for increased wheel-running behavior in house mice. *Behavior Genetics*, *28*, 227–237.
- Van Landuyt, L. M., Ekkekakis, P., Hall, E. E., & Petruzzello, S. J. (2000). Throwing the mountains into the lakes: on the perils of nomothetic conceptions of the exercise-affect relationship. *Journal of Sport and Exercise Psychology*, *22*, 208–234.
- Wagner, P. D. (2008). Commentary on viewpoint: perspective on the future use of genomics in exercise prescription. *Journal of Applied Physiology*, *104*, 1246.
- Williams, D. M., Dunsiger, S., Ciccolo, J. T., Lewis, B. A., Albrecht, A. E., & Marcus, B. H. (2008). Acute affective response to a moderate-intensity exercise stimulus predicts physical activity participation 6 and 12 months later. *Psychology of Sport and Exercise*, *9*, 231–245.
- Wipfli, B. M., Rethorst, C. D., & Landers, D. M. (2008). The anxiolytic effects of exercise: a meta-analysis of randomized trials and dose-response analysis. *Journal of Sport and Exercise Psychology*, *30*, 392–410.
- Wolfarth, B., Bray, M. S., Hagberg, J. M., Perusse, L., Rauramaa, R., Rivera, M. A., et al. (2005). The human gene map for performance and health-related fitness phenotypes: the 2004 update. *Medicine and Science in Sports and Exercise*, *37*, 881–903.
- Zubieta, J. K., Heitzeg, M. M., Smith, Y. R., Bueller, J. A., Xu, K., Xu, Y., et al. (2003). COMT *val¹⁵⁸met* genotype affects μ -opioid neurotransmitter responses to a pain stressor. *Science*, *299*, 1240–1243.

Panteleimon Ekkekakis*

Department of Kinesiology, Iowa State University, Ames, IA 50011, USA

* Tel.: +1 515 294 8766; fax: +1 515 294 8740.

E-mail address: ekkekaki@iastate.edu

4 September 2008