New Zealand Psychologist, 1979, *8*, 10-12.

If You Want to Know How Happy I am You'll Have to Ask Me

Robyn Irwin, Richard Kammann, and Graeme Dixon
University of Otago

The correlation between self-ratings of happiness and ratings provided by two flatmates in a sample of 78 young adults in Dunedin produced an *r* of .19 which is consistent with low *r*s observed in two previous studies. The present study also revealed an *r* of .40 between one's self-rating and the ratings of happiness assigned to flatmates, indicating a "projection" of own happiness onto others.

Most people imagine, and with some confidence, that they can accurately judge the happiness of friends and relatives. However two previous psychological studies which have carefully tested this assumption do not support it, nor does the present study.

In an older study on 195 American university students, Hartmann (1934) found an *r* of .34 between self-ratings of happiness and the mean rating of four student friends. In a much more recent study of 222 representative American adults, Andrews and Withey (1976, Chapter 6) report data from which we can derive an *r* of .33 between self-ratings of happiness and the mean rating of either two or three friends. Furthermore, Andrews and Withey found that ratings by friends failed to reflect differences among self-ratings of satisfaction with standard of living, the national government, leisure time, personal freedom, and housing, as well as the individual's life as a whole. The friends tended to view the subject as being more-or-less satisfied or dissatisfied with all these areas.

This is very similar to the halo effect or stereotyped perceptions which Campbell and Fiske (1959) found in a number of studies in which peers attempted to judge personality traits. One could typically make a better guess about the rating by knowing which friend or peer had made it, than by knowing which trait was being judged. Not surprisingly, accuracy was generally low. Considering those 29 comparisons in which self-ratings on a trait were compared with peer-ratings, the *r*s ranged from -.07 to .56 with a median *r* of .21.

We are not concerned here with personality traits, however, but with happiness which we view as a state of experience which is stable over two weeks and probably longer (Kammann, Christie, Irwin, & Dixon, 1979).

The present study is another comparison of self and peer-ratings of happiness which has three technical advantages over the Hartmann and the Andrews and Withey studies. First, we used an inventory of happiness which is considerably more reliable than their single item rating scales. Second, we did not allow subjects to nominate friends, which could have introduced sources of bias in others' ratings; instead, we obtained the co-operation of trios of flatmates in which each person rated himself or herself and both other members of the trio. Third, because each flatmate who rated others' happiness also rated his own happiness, we could determine if there was any tendency to "project" one's own happiness/unhappiness onto others.

Method

Subjects

The subjects were recruited by the first author through an informal acquaintance network in a search for flats which contained at least three people among whom there were no married couples or pairs of lovers. The subjects were 38 females and 40 males (N = 78) mostly living in the university area of Dunedin with a mean age of 22.1 years, and an *SD* of 4.7; forty percent were university students. Among the 26 trios were 21 mixed-sex groups, 3 all-female groups, and two all-male groups.

Measures

Each subject was given a packet containing three copies of the happiness inventory of 112 items described in Kammann *et al.*, (1979). The subject rated how often he or she had experienced each item, on a five-step rating scale with a time-set "over the past few weeks." The inventory also included a separate seven-step happiness rating scale from "extremely happy" to "fairly unhappy." The subject filled out one inventory on himself or herself, and the other two inventories on the two flatmates. The other-ratings had two different instructional sets: for one flatmate the subject was to fill out the questionnaire "as you think this person WILL fill it out" while for the other flatmate the instruction was, "as you think this person SHOULD fill it out if he were being totally honest with himself or herself." This provided a check on the possibility that there might be a difference in accuracy depending on the subject's rating task.

Results

Since none of the analyses revealed any differences as a result of the two different rating instructions (i.e., the other-will and other-should instructions) our results ignore this distinction and treat them simply as two other-ratings.

The ratings provided four different kinds of happiness relationships. We can explain these by designating one subject in each trio as Person A, and the two flatmates as B and C, although it is to be understood that the complete data collection allowed every person to be Person A.

Perceptual accuracy, the measure of primary interest here, was the relationship between A's self-rating (called A-A), and B's or C's rating of A. The B-A and C-A other-ratings were each correlated separately with the A-A ratings, and then the *r*s were averaged.

Projection was defined by the correlation between A's self-rating (A-A) and A's rating of the flatmates, A-B and A-C ratings. As with accuracy, these correlations were calculated separately for each other-rating and then averaged.

Pair similarity was the extent to which the self-ratings in a flat tended to be similar. In a trio there were three self-ratings which can be designated as A-A, B-B, and C-C, giving three possible pairs of self-ratings which could be correlated; all three pairs from each flat were entered into a single correlation across all flats.

Others' Agreement referred to the correlation between the two other-ratings on a given subject, that is the correlation between B-A and C-A ratings. This is an index of inter-judge reliability.

The results of these four analyses are presented in Table 1 for the subscales of the happiness inventory and for *net all* which is the overall inventory score. Parallel results are also reported for the seven-step happiness item. Because the inventory subscales merely confirm the net all scores, and because the seven-step happiness item is less reliable, we confine the discussion to net all scores.

Our estimate of perceptual accuracy (r = .19)[1] is actually lower than the value of .34 reported by Hartmann and of .33 derived from Andrews and Withey. Although our correlation is not significant, it is not our point that there is zero accuracy of perception, but that the accuracy is low. Accuracy of happiness judgments can be represented by the (N-weighted, Z-transformed) mean r of the three studies along with the associated 95 percent confidence limits (based on the total N which is 495) yielding a mean r = .31 with upper and lower bounds of .23 and .40. Although this procedure combines three studies which differ in method and subject samples, it provides a reference point for the accuracy of peer judgements of happiness.

Returning to the present data, we find a highly significant projection effect with r = .40 ($p < .001$) for net all scores. We conclude that one factor in the misperceptions of the happiness of others is the use of one's own happiness in making the judgment.[2]

Although similarity is known to be a strong factor in the selection of friends, it should not be possible to select flatmates on similarity of happiness if happiness is inaccurately perceived. This same result can be independently predicted on the argument that happiness is not a personality trait but is typically a slowly changing state. In any case, the net all r for pair similarity was a non-significant .13.

Table 1
Correlations (rs) on Four Types of Happiness Comparisons[a]

Measure	Perceptual Accuracy	Projection	Pair Similarity	Others' Agreement
Inventory				
Positive Sentences	.16	.43	.12	.32
Negative Sentences	.12	.39	.20	.32
Positive Adjectives	.19	.48	.10	.21
Negative Adjectives	.21	.40	.01	.25
Net Sentences	.14	.39	.16	.39
Net Adjectives	.21	.40	.05	.30
Net All	.19	.40	.13	.36
7-Step Happiness Item	.27	.29	.06	.30

[a] *With* N = 78, which applies to all correlations reported, correlations with $r > .22$ are significant at $p < .05$, and with $r > .28$ are significant at $p < .01$.

[1] The low r of .19 for perceptual accuracy is not appreciably based on the restriction of variability in other-ratings. The *SD* of net all for self-ratings was 1.01 against an *SD* of .83 for other-ratings (mean of two sets). The corresponding means of net all scores in a scale range from -4 to 4 were 1.10 and 1.16 for self and others' ratings, respectively.

[2] That the projection effect (r = .40) is not an artifact of pair similarity is shown by removing others' self-ratings (B-B and C-C ratings), giving a partial r = .39 for projection.

Other's agreement produced an r of .36 in terms of net all scores, and this is lower than our derivation of equivalent interjudge reliabilities in Hartmann's study (.50) or in the Andrews and Withey study (.51). This lower value also occurs for the seven-step happiness scale (r = .30) although this scale is very similar to the scales used by Hartmann and by Andrews and Withey.

The low interjudge reliability does not, however, explain the low accuracy of perception. With a reliability of .98 for self-ratings (Kammann, *et al.*, 1979) and an interjudge reliability of .36, the maximum possible r for self-peer ratings is .59 against an observed r of .19. The equivalent maximum r in Hartmann's data is .59 against an observed r of .34; the maximum r from Andrews and Withey is .58 against an observed r of .33. Although interjudge reliability sets a rather low ceiling on the self-other correlation, the observed rs are far below their ceilings indicating that ratings by others are measuring something different, such as own happiness (projection effect) or person stereotypes (halo effect).

We rejected the hypothesis that the low accuracy of happiness perception was due to a lack of acquaintance among the members of the trio. Virtually all of the flats had been together for at least the seven months of the academic year, and the majority had known each other much longer than that. Since we had collected information about length of acquaintanceship we were able to divide the sample into the more-acquainted pairs and less-acquainted pairs, which produced self-peer rs of .24 and .15 respectively, in terms of net all scores. While this is only a rough check, it does indicate that long acquaintance does not appreciably improve the perception of happiness.

Discussion

A plausible reason for why judgments of happiness are not accurate is that happiness is an *experience* and not a set of behaviours. One person may be happy in an energetic social way, while another may be happy in a quiet personal way which might be described as "contentment" or "tranquility". Similarly, one unhappy person may be socially energetic in a compulsive, anxious way, while another may be withdrawn and "depressed". Furthermore, social reinforcements may act to inhibit the free expression of unhappiness and dissatisfaction.

Given an absence of reliable behaviour to index a person's sense of well being, the friend or relative judging happiness may fall back on two extraneous sources of information: (a) his general positive-negative assessment of the person, expressing itself as a stereotype or halo effect across a number of traits, states, or attitudes: (b) his own level of happiness which biases his judgment of the other person, the "projection" effect.

What has been said so far about the perception of happiness may not apply with equal force to immediate moods. Kammann *et al.*, (1979) found that subjects could distinguish, in their self reports, between mood "right now" and happiness "over the past few weeks." Because happiness in this sense is an abstract judgment representing a characteristic level of experience, it should not include the greater extremes which can occur in short-term moods, and it may be that extreme moods are more expressive and more accurately perceived. This might help account for the apparent overconfidence people have in their ability to make judgments of "happiness", but this remains a conjecture.

A final point which must be left for further thought and future research is whether or not the inaccuracy of happiness perception is in any sense a "bad" thing, and whether or not people might be "helped" by learning how to communicate their happiness and sense of well being to each more effectively.

References

Andrews, F.M., & Withey, S.B. *Social indicators of well-being*. New York, London: Plenum Press, 1976.

Campbell, D.T., & Fiske, D.W. Convergent and discriminant validation by the multi-trait, multi-method matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

Hartmann, G.W. Personality traits associated with variations in happiness. *Journal of Abnormal Psychology*, 1934, *29*, 202-212.

Kammann, R, Christie, D.A., Irwin, R.M., & Dixon, G. Properties of an inventory to measure happiness (and psychological health). *New Zealand Psychologist*, 1979, *8*, 1 - 9.