Timing Effects of Listing Gratitude toward One's Parent(s) on Subjective Well-Being in Japanese Undergraduate Students

Futoshi Kobayashi

Author Note

I would like to express deep appreciation to those who participated in this study at Miyazaki International College, and Professors Jun Maeda, Haruko Aito, and Koji Watanabe (Miyazaki International College), who helped with data collection for this study. In addition, I would like to thank Professors Phillip Bennett and Julia Christmas (Miyazaki International College), who helped in the revision of this paper.

A different version of this article was presented at the 29th Annual Convention of the Association for Psychological Science, Boston, MA, May 25-28, 2017.

Please address correspondence regarding this article to Futoshi Kobayashi (fkobayas@miyazaki-mic.ac.jp), Miyazaki International College, 1405 Kano Kiyotake-cho Miyazaki-shi Miyazaki 889-1605 JAPAN.

Abstract

This study investigated the effectiveness of gratitude-listing interventions with two different timings, each with the same amount of practice (i.e., once a day for six days for a total of six times vs. once a week for five weeks for a total of six times) regarding subjective well-being. Seventy-five participants were randomly assigned to one of three conditions (i.e., listing three things for which they felt gratitude toward their parent(s) every day during six days for a total of six times, or once a week during five weeks for a total of six times, or no treatment). Results provided empirical evidence of effectiveness of the condensed gratitude listing practice (i.e., every day during six days for a total of six times) for subjective happiness and of effectiveness of the spaced-apart gratitude listing practice (i.e., once a week during five

weeks for a total of six times) for affect balance compared with the measurement-only control group. Further studies are needed to investigate this issue more thoroughly.

Keywords: gratitude, Japan, positive psychology, parent

Since Emmons and McCullough (2003) developed the original gratitude-listing interventions, many researchers have tried different variations of these interventions in positive psychological research (See Davis et al., 2016, for a review). In meta-analyzing five different studies, Davis et al. (2016) reported that gratitude-listing interventions demonstrated significant effects to improve well-being of participants when compared with measurement-only control groups.

However, there are still unknown factors that may influence the effectiveness of gratitude listing interventions. For example, Layous and Lyubomirsky (2014) pointed out the factors of timing and dosage. They used the metaphor of patients who are taking prescription drugs for their treatments. How often the patients take pills and how many they take makes a difference to the treatment effects. In the same vein, the effects of gratitude-listing interventions may fluctuate due to timing (i.e., frequency) and dosage (i.e., sum) of the gratitude listing practice.

Regarding the issue of timing and dosage, Dr. Robert Emmons, who is an authority in gratitude research, asserts the following:

Occasional gratitude journaling (e.g., twice weekly) boosts well-being more than the regular practice (e.g., every day) of counting blessings. Sometimes less is more. You avoid gratitude fatigue this way (Emmons, 2013, p. 12).

In addition, Lyubomirsky, Sheldon, and Schkade (2005) reported on a six-week gratitude listing intervention study. Lyubomirsky and her colleagues randomly made three groups that did gratitude listing (a) once a week, (b) three times a week, and (c) did nothing. They found that participants who did gratitude listing once a week, not three times a week, significantly increased their happiness over those who were in the control group. According to Emmons (2013) and Lyubomirsky, Sheldon, and Schkade (2005), lesser frequencies and lower dosages seem effective in some situations, but there are no clear guidelines regarding the issue of timing and dosage.

To my knowledge, there are no published experimental studies comparing the timing effects of the same amount of gratitude listing practice. Even though the total number of the pills that the patient takes is equal, the effect of the pill may be different if s/he takes three pills once a day or one pill every eight hours three times per day. In other words, the present study compared the timing effect for the same dosage of gratitude-listing interventions. Moreover, a certain type of gratitude, gratitude towards one's parent(s), was investigated in the present study because it has been emphasized in Japanese culture and history (Matsudaira, 1984; Oohata, 1971; Shintou, 1986). Previous research reported that daily listing of five things for which gratitude was felt toward one's parent(s) had improved the scores of subjective happiness and life satisfaction, however the daily events listing group, a control group with a matched activity, also improved the scores of life satisfaction and empathy via one-week intervention (Kobayashi, 2014). Additionally, other gratitude-listing intervention studies with Japanese samples did not exhibit significant intervention effects on well-being over control groups (Aikawa, Yada, & Yoshino, 2013; Otsuka, Hori, & Kawahito, 2012).

Three groups were randomly formed in this study. Group A, as a replication and a slight revision of gratitude listing of Kobayashi (2014), involved the processes of retrieving three things for which participants felt gratitude toward their parent(s), and required them to notice the three things by daily listing for six consecutive days (i.e., six entries). Group B was asked to retrieve three things for which they felt gratitude for their parent(s) and notice them by listing once a week for five weeks (i.e., six entries). Group C, as a control group, had no particular assignments for this study.

According to Ryan and Deci (2001), subjective well-being is defined with "three components: life satisfaction, the presence of positive mood, and the absence of negative

mood, together often summarized as happiness" (p. 144). In order to measure the well-being of the participants, three well-established subjective well-being measurements (i.e., subjective happiness, life satisfaction, & positive and negative affect) were used following the definition of Ryan and Deci (2001). All participants responded to these three measurements three times: just before the intervention started (Time 1), seven days later (Time 2), and five weeks and a day (i.e., 36 days) later (Time 3).

Previous studies that used similar one-week gratitude-listing intervention strategies demonstrated lasting effects toward subjective happiness in a one-month follow-up (Gander, Proyer, Ruch, & Wyss, 2013; Seligman, Steen, Park, & Peterson, 2005) and a composite score of life satisfaction and affect balance in a five-week follow-up (Watkins, Uhder, & Pichinevskiy, 2015) after the intervention was terminated.

This research was designed to be exploratory in its nature instead of testing particular hypotheses. I investigated how the subjective well-being of participants in Groups A, B, and C would change as a research question. Especially, I compared the improvements of each dependent variable for Group A from Time 1 to Time 2 and for Group B from Time 2 to Time 3 in order to check the effectiveness of the different timing schedules. In following the new statistics guidelines (Cumming, 2012), I investigated effect sizes and confidence intervals more thoroughly instead of searching for statistically significant results so that the results of the present study would be useful for future meta-analyses. When making statistical judgments on the results, I considered all the results of statistical significance levels, effect sizes, and confidence intervals, instead of relying on the statistical significance levels only.

Method

Participants

There were 25 participants in Group A, 25 participants in Group B, and 25 participants in Group C, for a total of 75. These 75 participants ranged in age from 18 to 47 years (M = 19.53, SD = 3.48). They were all undergraduate students (49 women and 26 men) who at the time were attending a liberal arts college in Miyazaki, Japan. They are all native speakers of Japanese.

Materials

Subjective Happiness. Shimai, Otake, Utsuki, Ikemi, and Lyubomirsky (2004) conducted a study with a sample of 302 Japanese undergraduate students in order to test the reliability and validity of the Japanese Subjective Happiness Scale (JSHS) that was created from the original Subjective Happiness Scale (Lyubomirsky & Lepper, 1999). The results indicated that the JSHS demonstrated appropriate factorial validity from factor analysis, and convergent and discriminate validity from significant correlations with relative scales. Regarding reliability, there was also an indication of sound internal consistency ($\alpha = .82$) and appropriate five-week test-retest reliability (r = .86). The JSHS has four items and the participants rate these items on a seven-point Likert-type scale and the sum values indicate their global subjective happiness. Kobayashi (2013) reported that it demonstrated appropriate construct validity by significant correlations with relative scales and sound internal consistency ($\alpha = .83$).

Positive and Negative Affect. In order to create a handy measurement to assess the emotional states of participants, Sato and Yasuda (2001) conducted two different studies with college-age Japanese samples. As a result, a Japanese version of the Positive and Negative

Affect Schedule (PANAS) was made from the original PANAS (Watson, Clark, & Tellegen, 1988). The Japanese PANAS has eight adjectives to measure positive affect ($\alpha = .90$) and another eight adjectives to measure negative affect ($\alpha = .91$). In this scale, participants indicate their emotional state by rating each adjective from 1 (*does not apply to me at all*) to 6 (*applies to me greatly*) on a Likert-type scale. Additionally, Kobayashi (2013) reported that the Japanese PANAS demonstrated sound internal consistency for positive affect ($\alpha = .87$) and negative affect ($\alpha = .88$) and appropriate construct validity.

Life Satisfaction. In order to create the Japanese version of the Satisfaction With Life Scale (SWLS) (Diener, Emmons, Larsen, & Griffin, 1985), Sumino (1994) conducted five different studies with various samples of Japanese university students and middle-age adults. The Japanese version of SWLS has five items that are intended to measure cognitive aspects of participants' subjective well-being via ratings that range from 1 (*strongly disagree*) to 7 (*strongly agree*) on a Likert-type scale. Regarding construct validity, it exhibited significant correlations with five relative measurements. It also showed sound reliability by acceptable internal consistency in an undergraduate sample ($\alpha = .84$) and a middle-age adult sample ($\alpha = .90$) and r = .80 for four-week test-retest reliability. Recently, it also demonstrated acceptable internal consistency ($\alpha = .83$) and appropriate construct validity with significant correlations with relative scales (Kobayashi, 2013).

Procedure

First of all, the Institutional Review Board and the Dean of the School of International Liberal Arts where the study took place granted me ethical approval to conduct this study. I visited several classes in the college where I work and explained the study to the undergraduate students. Interested students visited my office to read the description of the study and their rights and rewards for participation. Many of them decided to participate in this study and provided informed consent. The participants randomly picked their own identification number cards (e.g., "B15" representative of the 15th participant of Group B) from a bag and used them in order to assure their anonymity whenever they answered the surveys and wrote their gratitude listings. All the participants anonymously responded to the three measurements of life satisfaction, subjective happiness, and positive and negative affect. Then, notebooks were given to all the participants of Groups A and B, and they retrieved and listed three things for which they felt gratitude toward their parent(s) for approximately five to ten minutes every night for Group A (or once a week for Group B) before going to bed. In order to avoid gratitude fatigue, participants of Groups A and B were asked to retrieve three things for which they felt gratitude toward their parent(s) when they were ages 3 to 6 on the first entry, ages 7 to 9 on the second entry, ages 10 to 12 on the third entry, ages 13 to 15 on the fourth entry, ages 16 to 18 on the fifth entry, and in college on the sixth entry. Seven days later, all participants returned to my office, answered the same surveys for the second time, and the participants of Group A returned their notebooks. Five weeks and a day later, all participants returned to my office and answered the same surveys for the third time and the participants of Group B returned their notebooks. Then, each participant who answered the surveys the full three times received 1,000 Japanese yen (approximately US \$10 in August 2016) as a financial reward.

Results

A 3 (between subjects: treatment group) X 3 (within subjects: time of assessment) Multivariate Analysis of Variance (MANOVA) was conducted toward three dependent variables: subjective happiness, life satisfaction, and affect balance. The score of affect balance was made by subtracting the total number of negative affect from that of positive

affect. The results revealed significant multivariate effects across the interaction between group and time, V = .311, F(12, 136) = 2.089, p = .021, $\eta_p^2 = .156$. However, no significant multivariate effects were found on time, V = .080, F(6, 67) = .973, p = .450, $\eta_p^2 = .080$ and group, V = .031, F(6, 142) = .377, p = .893, $\eta_p^2 = .016$. Consequently, a 3 (between subjects: treatment group) X 3 (within subjects: time of assessment) Analysis of Variance (ANOVA) was conducted toward each dependent variable. Mauchly's tests indicated that the assumption of sphericity had been violated regarding subjective happiness, $\chi^{2}{}^{(2)} = 12.76$, p = .002 and affect valance, $\chi^2{}^{(2)} = 7.06$, p = .029. Their estimates of sphericity were greater than .75 and thus the Huynh-Feldt estimates were reported in the present study (See Field, 2013, p. 548). The results indicated significant interaction effects on affect balance, F(3.849, 138.6) = 4.489, p = .002, $\eta_p^2 = .111$. There were marginally significant interaction effects on subjective happiness, F(3.609, 129.9) = 2.169, p = .083, $\eta_p^2 = .057$, and there were no significant interaction effects on life satisfaction, F(4, 144) = 1.574, p = .184, $\eta_p^2 = .042$.

Regarding subjective happiness, post-hoc tests with a Bonferroni adjustment revealed that Group A displayed a marginally significant increase in their scores from Time 1 to Time 2, t(24) = 2.14, p = .108, mean difference = 0.920, 95% CI = [-0.134, 1.974], Cohen's d = 0.428, and Group B increased their scores from Time 2 to Time 3, t(24) = 1.49, p = .420, mean difference = 0.720, 95% CI = [-0.462, 1.902], Cohen's d = 0.299. See Table 1 and Figure 1.

Regarding affect balance, post-hoc tests with a Bonferroni adjustment revealed that Group A's scores were significantly decreased from Time 2 to Time 3, t(24) = 2.60, p = .034, mean difference = 5.640, 95% CI = [0.327, 10.953], Cohen's d = 0.520, and Group C's scores also decreased significantly from Time 1 to Time 3, t(24) = 2.63, p = .031, mean difference = 6.040, 95% CI = [0.413, 11.667], Cohen's d = 0.526. Group A's scores increased from Time 1 to Time 2, t(24) = 1.31, p = .587, mean difference = 2.280, 95% CI = [-1.998, 6.558], Cohen's d = 0.261, and Group B's scores increased with marginal significance from Time 2 to Time 3, t(24) = 2.38, p = .060, mean difference = 5.160, 95% CI = [-0.153, 10.473], Cohen's d = 0.476. See Table 1 and Figure 2.

Regarding life satisfaction, post-hoc tests with a Bonferroni adjustment revealed that Group A's scores increased from Time 1 to Time 2, t(24) = 0.82, p = 1.00, mean difference = 0.600, 95% CI = [-1.193, 2.393], Cohen's d = 0.164, and Group B's scores increased from Time 2 to Time 3, t(24) = 1.76, p = .247, mean difference = 1.200, 95% CI = [-0.470, 2.870], Cohen's d = 0.352. See Table 1 and Figure 3.

Discussion

In following the new statistics guidelines (Cumming, 2012), I considered all the results of statistical significance levels, effect sizes, and confidence intervals and would like to discuss four major findings of this study.

First, the subjective happiness of Group A seemed to increase from Time 1 to Time 2 but did not decrease at Time 3. This finding indicates that the condensed gratitude listing practice (i.e., every day during six days for a total of six times) might be effective to increase one's subjective happiness and such an effect might continue for four weeks after the termination of such a gratitude practice. Such an increase from Time 2 to Time 3 of Group B who conducted the spaced-apart gratitude listing practice (i.e., once a week during five weeks for a total of six times) was negligible.

Second, the affect balance of Group B seemed to increase from Time 2 to Time 3. This finding indicates that the spaced-apart gratitude listing practice (i.e., once a week during five weeks for a total of six times) might be effective to increase one's affect balance. Such an increase from Time 1 to Time 2 of Group A, who conducted the condensed gratitude listing practice (i.e., every day during six days for a total of six times) was negligible.

Third, the affect balance of Group C seemed to decrease from Time 1 to Time 3. Additionally, those who were in Group A did not do anything from Time 2 to Time 3 and this seemed to decrease their affect balance. These results indicate that one's affect balance might be decreased when one does not do any gratitude activities. Additionally, I suspect that the particular semester schedule in which the study was conducted might have influenced the results. Most of the participants started their assignments at the beginning of the semester when there was not much demanding school work. As the semester went by, the amount of school work increased and it might have influenced their level of affect balance.

Fourth, both the condensed and the spaced-apart gratitude listing practices seemed to make no significant effects on life satisfaction. It is intriguing because previous similar studies (Kobayashi, 2014, 2015) found the life satisfaction scale which was used in this study was useful for measuring the effects of gratitude-related practices. This might have resulted from idiosyncrasies of the particular sample.

There are some weaknesses in this study. First of all, this study used a convenience sample from a particular institution. Second, the sample size was small. Therefore, all results of this study might be the products of idiosyncrasies in the particular sample. Third, the results were based on self-reporting by the participants, therefore, there is danger of self-serving bias. Fourth, many participants of this study, who are native Japanese speakers, might be qualitatively different from a general Japanese population because many of them study liberal arts in English and they live their college life in an atypical, nearly English-only environment.

Although this study has certain weaknesses, it could provide empirical evidence of the effectiveness of the condensed gratitude listing practice (i.e., every day during six days for a total of six times) for subjective happiness and of effectiveness of the spaced-apart gratitude listing practice (i.e., once a week during five weeks for a total of six times) for affect balance

compared to the measurement-only control group. Further studies, especially meta-analyses, are needed to investigate this issue more thoroughly.

References

- Aikawa, A., Yada, S., & Yoshino, Y. (2013). Kanshya wo kazoerukoto ga shukanteki uerubiinngu ni oyobosu kouka ni tuite no kainyu jikken [An interventional experiment on the effect of counting gratitude on subjective well-being]. *Tokyougakugeidaigakukiyou sougoukyoikukagakukei I, 64*, 125-138.
- Cumming, G. (2012). Understanding the new statistics: Effect size, confidence interval, and meta-analysis. New York, NY: Routledge.
- Davis, D. E., Choe, E., Meyers, J., Wade, N., Varjas, K., Gifford, A., Quinn, A., Hook, J. N.,
 Van Tongeren, D. R., Griffin, B. J., & Worthington, E. L., Jr. (2016). Thankful for
 little things: A meta-analysis of gratitude interventions. *Journal of Counseling Psychology*, 63, 20-31. doi:10.1037/cou0000107
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49, 71-75. doi:10.1207/s15327752jpa4901_13
- Emmons, R. A. (2013). *Gratitude works! A 21-day program for creating emotional prosperity*. San Francisco: Jossey-Bass.
- Emmons, R. A., & McCullough, M. E. (2003). Counting blessings versus burdens: An empirical investigation of gratitude and subjective wellbeing in daily life. *Journal of Personality and Social Psychology*, 84, 377-389. doi:10.1037/0022-3514.84.2.377
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics and sex and drugs and rock 'n' roll*. (4th ed.). Thousand Oaks, CA: Sage.
- Gander, F., Proyer, R. T., Ruch, W., & Wyss, T. (2013). Strength-based positive interventions: Further evidence for their potential in enhancing well-being and alleviating depression. *Journal of Happiness Studies*, *14*, 1241-1259. doi:10.1007/s10902-012-9380-0

- Kobayashi, F. (2013). Development and validation of the Gratitude Questionnaire (GQ) in Japanese undergraduate students. *Comparative Culture, 18,* 1-18.
- Kobayashi, F. (2014). Daily gratitude towards one's parents and subjective well-being of Japanese undergraduate students. *Comparative Culture, 19*, 1-17.
- Kobayashi, F. (2015). The effect of recalling relationships with parents on self-acceptance and life satisfaction. *Comparative Culture*, *20*, 1-15.
- Layous, K., & Lyubomirsky, S. (2014). The how, why, what, when, and who of happiness: Mechanisms underlying the success of positive interventions. In J. Gruber & J.
 Moscowitz (Eds.), *Positive emotion: Integrating the light sides and dark sides* (pp. 473-495). New York: Oxford University Press.
- Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46*, 137-155. doi:10.1023/A:1006824100041
- Lyubomirsky, S., Sheldon, K. M., & Schkade, D. (2005). Pursing happiness: The architecture of sustainable change. *Review of General Psychology*, *9*, 111-131. doi:10.1037/1089-2680.9.2.111
- Matsudaira, K. (1984). *Nisshinkan doujikun gendaigoyaku* [Ethics textbook at Nisshinkan in contemporary Japanese language]. Tokyo, Japan: Sanshintosho.

Oohata, S. (1971). Doutoku no rekishi [History of ethics]. Tokyo, Japan: Yuzankaku.

- Otsuka, Y., Hori, M., & Kawahito, J. (2012). Improving well-being with a gratitude exercise in Japanese workers: A randomized controlled trial. *International Journal of Psychology and Counseling*, *4*, 86-91. doi:10.5897/IJPC11.031
- Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology*, 52, 141-166. doi:10.1146/annurev.psych.52.1.141

- Sato, A., & Yasuda, A. (2001). Nihongo ban PANAS no sakusei [Development of the Japanese version of Positive and Negative Affect Schedule (PANAS) scales]. Japanese Journal of Personality, 9, 138-139.
- Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: Empirical validation of interventions. *American Psychologist*, 60, 410-421. doi:10.1037/0003-066X.60.5.410
- Shimai, S., Otake, K., Utsuki, N., Ikemi, A., & Lyubomirsky, S. (2004). Nihonban syukanteki koufukukan shakudo (Subjective Happiness Scale: SHS) no shinraisei to datousei no kentou [Development of a Japanese version of the Subjective Happiness Scale (SHS), and examination of its validity and reliability]. *Japanese Journal of Public Health,* 51, 845-853.
- Shintou, K. (1986). *Shiryou "tennoushouchoku" senshuu* [Reference materials for selected Emperors' imperial rescripts]. Tokyo, Japan: Nippon gitsugyou shuupansha.
- Sumino, Z. (1994). Jinsei ni taisuru manzoku syakudo (the Satisfaction With Life Scale [SWLS]) nihonban sakusei no kokoromi [Development of the Japanese version of the Satisfaction With Life Scale]. Proceedings of the 36th Annual Meeting of the Japanese Association of Educational Psychology, 36, 192.
- Watkins, P. C., Uhder, J., & Pichinevskiy, S. (2015). Grateful recounting enhances subjective well-being: The importance of grateful processing. *Journal of Positive Psychology*, *10*, 91-98. doi:10.1080/17439760.2014.927909
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality* and Social Psychology, 54, 1063-1070. doi:10.1037/0022-3514.54.6.1063

Table 1Means, Standard Errors, and 95% Confidence Intervals by Condition and Time of Assessment

			Time of Assessment					
DV	Condition	п	Time 1 (<i>SE</i>)	95% CI	Time 2 (<i>SE</i>)	95% CI	Time 3 (<i>SE</i>)	95% CI
Subjective Happiness	Group A	25	19.12 (0.764)	[17.60, 20.64]	20.04 (0.766)	[18.51, 21.57]	20.12 (0.728)	[18.67, 21.57]
	Group B	25	18.68 (0.764)	[17.16, 20.20]	18.32 (0.766)	[16.79, 19.85]	19.04 (0.728)	[17.59, 20.49]
	Group C	25	19.60 (0.764)	[18.08, 21.12]	18.92 (0.766)	[17.39, 20.45]	18.84 (0.728)	[17.39, 20.29]
Affect Balance	Group A	25	8.60 (1.78)	[5.05, 12.15]	10.88 (2.28)	[6.34, 15.42]	5.24 (2.33)	[0.60, 9.88]
	Group B	25	7.08 (1.78)	[3.53, 10.63]	4.04 (2.28)	[50, 8.58]	9.20 (2.33)	[4.56, 13.84]
	Group C	25	11.88 (1.78)	[8.33, 15.43]	8.80 (2.28)	[4.26, 13.34]	5.84 (2.33)	[1.20, 10.48]
Life Satisfaction	Group A	25	19.76 (1.05)	[17.67, 21.85]	20.36 (1.22)	[17.92, 22.80]	20.44 (1.25)	[17.96, 22.92]
	Group B	25	18.96 (1.05)	[16.87, 21.05]	18.96 (1.22)	[16.52, 21.40]	20.16 (1.25)	[17.68, 22.64]
	Group C	25	20.48 (1.05)	[18.39, 22.57]	20.64 (1.22)	[18.20, 23.08]	19.64 (1.25)	[17.16, 22.12]

Note. SE = standard error, CI = confidence interval, DV = dependent variable.



Figure 1. Changes of subjective happiness at three time periods: Time 1 (Pretest), Time 2 (7 days later) and Time 3 (36 days later). Error bars indicate 95% confidence intervals.



Figure 2. Changes of affect balance at three time periods: Time 1 (Pretest), Time 2 (7 days later) and Time 3 (36 days later). Error bars indicate 95% confidence intervals.



Figure 3. Changes of life satisfaction at three time periods: Time 1 (Pretest), Time 2 (7 days later) and Time 3 (36 days later). Error bars indicate 95% confidence intervals.