Revised Description and Measurement of Ego Development in Early Adolescence:: An Artifact of the Written Procedure?
P. Michiel Westenberg, Suzanne D. van Strien and Martine J. Drewes
*The Journal of Early Adolescence* 2001; 21; 470
DOI: 10.1177/0272431601021004005

The online version of this article can be found at: http://jea.sagepub.com/cgi/content/abstract/21/4/470
Revised Description and Measurement of Ego Development in Early Adolescence: An Artifact of the Written Procedure?

P. Michiel Westenberg
Suzanne D. van Strien
Martine J. Drewes
Leiden University, the Netherlands

The applicability of the Loevinger conception of ego development was explored for the period of early adolescence. Recent modifications of the earliest ego levels are summarized, and a new version of the measure is presented: the Sentence Completion Test for Children and Youth (SCT-Y). Questions were considered about the reliability and validity of this new instrument. The main question addressed was whether an oral administration would yield essentially different responses and significantly differing ego-level scores as compared to the standard written administration in an early adolescent sample (9.5 through 15.5 years of age; N = 120). The results indicated that an oral administration of the SCT-Y did not yield essentially different responses or significantly differing ego-level scores, regardless of the respondent’s age, gender, reading/writing skills, and preference for either presentation mode. The findings indicate that the revision of the ego development construct and measure for (early) adolescence does not seem to be an artifact of the written procedure.

Even though the early adolescent period marks psychosocial developments common to all, it is obvious that young adolescents differ in many ways. For example, early adolescence is viewed as a normative period of increasing independence (i.e., emotional autonomy and individuation from parents; see Steinberg, 1999), yet at the same time some young adolescents are much more independent than are others. The question is whether the latter could be attributed to the former, that is, whether individual differences in the psychosocial realm (e.g., independence) might be related to developmental differences.

According to the Loevinger (1976, 1997) conception of ego development, individuals differ in the speed, timing, and extent of development, leading to significant differences in developmental maturity at any given age—a pre-
sumption clearly supported by the findings of longitudinal studies (e.g., Gfeller, 1986; Hauser, Powers, & Noam, 1991; Westenberg & Gjerde, 1999). Differential maturity would contribute to individual differences within any age cohort. In other words, young adolescents differ in many ways, partly because they differ in terms of the developmental level attained (e.g., Westenberg & Block, 1993). The possibility that individual differences might be due to developmental differences is overlooked frequently by researchers and practitioners working with (young) adolescents.

The Loevinger Conception of Ego Development

The Loevinger (1976, 1993) conception of ego development concerns the individual’s central frame of reference, or master trait, regarding the psychosocial realm. McAdams (1998) likens ego to the William James “I”—the self as subject, whereas dispositional traits, personal concerns, and life narratives reflect the “Me”—the self as object. “The ego’s relation to [these] three levels of personality is that of the I to the Me…. Loevinger’s ego should function as the master orchestrator of traits, concerns, and narrations” (McAdams, 1998, p. 35). Psychoanalysts also have postulated an organizing or synthesizing function of the ego in addition to the other ego functions (e.g., defense mechanisms). Loevinger (1976), however, has argued that the synthesizing function is not one among many ego functions but denotes what ego is: “The striving to master, to integrate, to make sense of experience is not one ego function among many but the essence of the ego” (p. 59).

The continuum of ego development consists of three strands—impulse control, interpersonal style, and conscious preoccupations—and is marked by nine qualitatively different milestones (see Table 1 for a brief outline). The low end of the scale is marked by the Impulsive ego level. Loevinger (1997) described individuals at the Impulsive ego level as easily yielding to their aggressive and sexual impulses and dependent on others for impulse control. They are oppositional and defiant and view rules as arbitrary and punishment as retaliatory. Impulsive individuals display an egocentric interpersonal style and expect others to cater to their needs. At the next, Self-Protective, ego level, control of self and others is viewed as crucial to further the person’s own interests. An opportunistic morality is coupled with a manipulative and exploitive attitude toward others. Rules are understood but manipulated. The overarching rule of Self-Protective individuals is to stay out of trouble and not be caught.

In contrast to the egocentric perspective characteristic of the Impulsive and Self-Protective individuals, persons at the Conformist ego level adopt a
sociocentric perspective. They are attuned to the needs, expectations, and opinions of others. Approval is valued, disapproval is feared. Everyone is, or ought to be similar, just as rules of conduct and appearance apply to everyone. The rules of the individual’s social group are accepted just because they are the rules. Conformity should not be confused with conventionality: The person might adhere rigidly to nonconventional standards. The next, Self-Aware, ego level is characterized by the awareness of personal feelings and thoughts both in self and others. The examination of inner life is accompanied by a sense of being different from others. Exceptions to rules are allowable, deviant behavior and opinions are tolerated. A good relationship is defined by the sharing of innermost feelings and thoughts. Empirical studies have indicated that the Self-Aware ego level is very rare in early adolescence, and the higher ego levels (see Table 1) are not at all relevant to young adolescents (e.g., Avery & Ryan, 1988; Cohn, 1998; Gfellner, 1986; Westenberg & Block, 1993).

Level of ego development is measured by means of the Washington University Sentence Completion Test (WUSCT) (Loevinger, 1985, 1998). The WUSCT yields 36 responses that are rated by an equal number of very

### TABLE 1: Some Characteristics of Ego Development Levels

<table>
<thead>
<tr>
<th>Ego Level</th>
<th>Impulse Control</th>
<th>Interpersonal Mode</th>
<th>Conscious Preoccupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2. Impulsive</td>
<td>Impulsive</td>
<td>Egocentric, dependent</td>
<td>Bodily feelings</td>
</tr>
<tr>
<td>E3. Self-Protective</td>
<td>Opportunistic</td>
<td>Manipulative, wary</td>
<td>Trouble, control</td>
</tr>
<tr>
<td>E4. Conformist</td>
<td>Respect for rules</td>
<td>Cooperative, loyal</td>
<td>Appearances, behavior</td>
</tr>
<tr>
<td>E5. Self-Aware</td>
<td>Exceptions allowable</td>
<td>Helpful, self-aware</td>
<td>Feelings, problems, adjustment</td>
</tr>
<tr>
<td>E6. Conscientious</td>
<td>Self-evaluated standards, self-critical</td>
<td>Intense, responsible</td>
<td>Motives, traits, achievements</td>
</tr>
<tr>
<td>E7. Individualistic</td>
<td>Tolerant</td>
<td>Mutual</td>
<td>Individuality, development, roles</td>
</tr>
<tr>
<td>E8. Autonomous</td>
<td>Coping with conflict</td>
<td>Interdependent</td>
<td>Self-fulfillment, psychological causation</td>
</tr>
</tbody>
</table>

detailed scoring manuals. The cumulative frequency of the 36-item ratings is converted to a total protocol rating by means of a set of objective scoring rules. Many studies have supported the construct validity of the WUSCT (see Carlson & Westenberg, 1998; Hauser, 1976; Loevinger, 1979; Manners & Durkin, 2000; Westenberg, Blasi, & Cohn, 1998). The presumption that differential maturity is related to individual differences within age cohorts has been demonstrated in numerous research studies that have addressed a wide variety of issues (e.g., John, Pals, & Westenberg, 1998; Pals & John, 1998).


Despite the impressive construct validity of the WUSCT as a measure of ego maturity, two questions surfaced as far as ego development in the early adolescent period is concerned: (a) Is a description of ego development based on adults also applicable to ego development in the early adolescent period? and (b) Is the sentence completion procedure used with adults also appropriate for measuring ego development in young adolescents?

Measuring and Describing Ego Development in Children and Adolescents

As noted previously, the Impulsive through Conformist ego levels are of particular relevance to late childhood and early adolescence. Yet the descriptions of those levels were based primarily on research with respondents who were older than 18 years of age (see Loevinger, 1985, 1998; Loevinger & Wessler, 1970). Likewise, the WUSCT was constructed for measuring ego level in (young) adults (Loevinger, 1993, 1998). On the basis of conceptual analyses and comparisons with other developmental models, however,
Loevinger (1976) has argued that the sequence of adult ego levels reflects developments occurring in childhood and adolescence. However, the empirical success of the measure obscured the fundamental question as to whether the stage descriptions and the measurement instrument would be adequate for describing and measuring ego level in children and youth. Are theoretical generalizations about the “lower” ego levels, as derived from adults, conceptually and empirically similar to the “early” ego levels of children and youth moving through those stages during a normal developmental course?

On the basis of a large-scale investigation of ego development in more than 2,500 children and youth, Westenberg, Treffers, and Drewes (1998) concluded that the WUSCT test protocol and scoring manual had to be modified in several ways to be adequate for measuring ego development of adolescents and children older than 8 years of age: the Sentence Completion Test for Children and Youth (SCT-Y)\(^1\) (see Loevinger, 1998). Moreover, the feedback loop between the scoring categories in the manual, on one hand, and the descriptions of the ego levels on the other hand (i.e., microvalidation; see Loevinger, 1993), resulted in changes in the descriptions of the earliest ego levels to be adequate for children and adolescents (Westenberg, Jonckheer, Treffers, & Drewes, 1998).

A brief account of the differences with the Loevinger description are provided here (for a full account of the similarities and differences with the Loevinger description of these ego levels, see Westenberg, Jonckheer, et al., 1998). The Impulsive child or adolescent did not display just aggressive or sexual impulses but also displayed empathic impulses and prosocial attitudes. The Impulsive child is preoccupied with aggression, but mostly in terms of being the victim and not in terms of being the perpetrator, and displays a general sense of vulnerability. Moreover, the oppositional and defiant attitude attributed to the Impulsive adult was not typical of the Impulsive child. Impulsive children appreciate concrete rules in specific situations, but due to their impulsivity they require supervision and regular reminders.

The research on children and adolescents also yielded a more balanced picture of the Self-Protective level. Consistent with the Loevinger description, control is of paramount importance, but the need for control is not displayed only through a manipulative and exploitive attitude toward others. In contrast, the Self-Protective child or adolescent prefers self-focused forms of control (e.g., a general denial of hurt feelings and disappointments), displays a self-reliant attitude, and has a live-and-let-live philosophy of life. Emotional and behavioral autonomy is valued more than controlling others. In contrast with the Loevinger description of Impulsive and Self-Protective
adults, children, and youth at these levels report positive interactions and appreciative relations with parents and friends.

Finally, the description of the Conformist level also was adjusted to be appropriate for children and adolescents. Central in the Loevinger description of this level is the mindless conformity to externally imposed and concrete rules of the social group of which the individual happens to be part. The Conformist child or adolescent, however, appears to adhere to a few ground rules, such as equality and reciprocity. He or she is likely to oppose specific rules that are inconsistent with those ground rules and will be inclined to shy away from nonegalitarian groups.

The differences were marginal for the Self-Aware level and beyond, probably due to the age overlap of participants at these levels in the Loevinger samples and the normative samples (Westenberg, Jonckheer, et al., 1998). The differences with the Loevinger model for the first three ego levels are attributed to the younger age of participants in the normative samples and to the greater representation of the lowest ego levels. Loevinger (1998) realized that by using adult samples, “One cannot extrapolate very far to ego levels that are underrepresented” (p. 24). On the basis of many low ego-level children and adolescents, however, “A more positive image of the earliest stages emerged” (Loevinger, Carlson, Westenberg, & Lasker, 1998, p. 52).

Many questions can be asked about the reliability and validity of the newly constructed SCT-Y, and some of those questions have been addressed already. The SCT-Y does very well in terms of the usual indices of reliability, showing high interrater agreement, internal consistency, and split-half and test-retest stability (Westenberg et al., 2000). In addition, Drewes and Westenberg (2001) demonstrated the robustness of the SCT-Y with respect to varying test instructions. The effects of two differing test instructions were studied with a split-half within-participants design. Respondents were asked to complete the first half of the SCT-Y under the standard instructions (“Complete the following sentences in any way that you wish”) and were asked to complete the second half under two modified instructions: (a) an instruction irrelevant to the ego development construct (a social desirability or “be good” instruction: “Complete the following sentences in such a way that you’ll make a very good impression on others”), and (b) an instruction highly relevant to the ego development construct (a high ego level or “be mature” instruction: “Complete the following sentences in as adult and mature a manner as you can”). As was expected, the “be good” instruction did not have an effect on ego-level scores, whereas the “be mature” instruction had a significant but small effect (an average raise of .19 ego level). The higher scores under the “be mature” instructions were interpreted to reflect the respondents’ optimal ego
level (best functioning), whereas the ego-level score under the standard instructions reflects their functional ego level (normal functioning).

The findings reported by Drewes and Westenberg (2001) are a first indication of the construct validity of the SCT-Y. The SCT-Y was not susceptible to an instruction that did not convey information relevant to the ego development construct (i.e., the “be good” instruction). In contrast, the SCT-Y was mildly susceptible to an instruction that was highly relevant to the construct of ego maturity (i.e., the “be mature” instruction). Another indication of the construct validity of the SCT-Y was provided by the findings of research on the relation between ego development and anxiety disorders in a population of children and adolescents referred to an outpatient psychiatric clinic (Westenberg et al., 1999). The two most prevalent and debilitating anxiety disorders in children and adolescents were related empirically to conceptually equivalent ego levels: (a) Separation anxiety disorder was empirically related to the Impulsive level (both share a number of key concepts, such as vulnerability and dependency), and (b) overanxious disorder was related to the Conformist level (both share the concern about the adequateness of behavior, appearance, or performance, and assurance seeking). Those relations were controlled statistically for socioeconomic status and IQ.

Those findings attest to the construct validity of the SCT-Y, but several additional issues still need to be addressed. A particularly pressing question is whether the revised scoring manual and new descriptions of the earliest ego levels might be an artifact of the written procedure used to study ego development in late childhood and adolescence. Would an oral mode of presentation be a more appropriate procedure for measuring ego development in late childhood and early adolescence? Would an oral procedure yield different, perhaps more elaborated responses, or would it lead only to differing ego level scores?

Written as Compared to Oral Administration of the SCT-Y

The standard administration procedure for the SCT-Y, as for the WUSCT, requires participants to write down their own completions, and the administrator is instructed to grant respondents sufficient privacy to do so. This procedure is consistent with the concept of ego development because it allows individuals to respond in their own way and pace and to display their own frame of reference without being guided by the administrator’s presence or reactions. In other words, the written procedure minimizes the demand characteristics of the test setting. Moreover, the written procedure is the most efficient way to collect SCT data (e.g., it allows for group testing). Yet the ques-
tion remains whether respondents, particularly children and young adolescents, would provide different types of responses and would obtain differing ego-level scores if the SCT-Y were to be administered orally. Some respondents might not be accustomed to expressing their thoughts and feelings in writing and/or might not have developed sufficient writing skills to do so. They might benefit from oral testing. In contrast, the written procedure generally would allow for more time to think about a response. More thoughtful and carefully crafted responses might lead to higher ego-level scores for the written mode.

To date, no study has been designed for investigation of an oral as compared to a written administration of the SCT-Y, but three such studies were conducted for the WUSCT (Hansell, Sparacino, Ronchi, & Strodtbeck, 1985; McGammon, 1981; Streich & Swensen, 1985). All three studies used the same design: One-half of the items were presented in the standard written mode, the other half were presented in the oral mode (counterbalancing the order of the presentation mode). Streich and Swensen (1985) studied three age groups: college age, community residents, and retired people ($n = 64$ in each group). They did not find a main effect for mode of presentation and did not find an interaction effect for age group by presentation mode. A significant interaction effect was observed for gender: “Women tended to score higher on the written half” (Streich & Swensen, 1985, p. 288) ($p < .05$). Hansell et al. (1985) studied the differences for written and telephone administration of the WUSCT in a mostly adult sample from a rural community ($N = 221$). The telephone ego-level scores generally were lower than the written scores ($p < .001$), and this effect was attributed to the greater time to reflect on and modify written responses. The difference between telephone and written scores was pronounced, particularly for women.

The only study in which adolescents were included was conducted by McGammon (1981), who studied a group of 6th and a group of 10th graders ($n = 40$ in each grade). The two presentation modes did not yield any ego-level differences for the boys, but there was a significant difference for the girls. For girls, the written mode yielded higher scores than the oral mode ($p < .01$ in both age groups). The interaction effect for age group was not significant: The 6th graders did not benefit more from either presentation than did the 10th graders. McGammon concluded that in the population sampled, “Starting at least with the sixth grade, boys and girls have developed sufficient writing skills to make oral testing unwarranted” (pp. 234-235).

In summary, the findings of those studies indicate that an oral mode of presentation of the WUSCT does not support the average adolescent or adult respondent to obtain higher ego-level scores. In fact, an unanticipated but consistent finding was that females obtained lower average scores with the
oral mode; they appeared to benefit from the written mode. That un-
anticipated finding needs to be studied more thoroughly. Moreover, a written/
oral study has not been conducted yet with respondents below the sixth grade.

RESEARCH AIMS

The general aim for the current study was to investigate possible differ-
ences between oral and written presentation modes for the newly constructed
SCT-Y for a sample of early adolescents (9.5 through 15.5 years of age,
Grades 4 through 9). Would oral responses be any different from written
responses, and would that difference lead to differing ego-level scores? Such
a study had not been conducted yet with the SCT-Y and as yet had not been
conducted with fourth- and fifth-grade students using the WUSCT. Below
the sixth grade, reading and writing skills might not have developed suffi-
ciently for a written administration of the SCT-Y; hence, the youngest partici-
pants might benefit from an oral presentation. This study was the first one
through which to take an explicit look at the effect of individual differences in
reading/writing skills in the youngest group (fourth- through sixth-grade stu-
dents) and to study the possible effect of a personal preference for either pre-
sentation mode. Interaction effects for gender were also studied systemati-
cally. This study was also the first for the study of the utility and reliability of
the measure under both presentation modes. The manual for the SCT-Y was
developed on the basis of written responses and might not be adequate for
classifying oral responses.

METHOD

Participants

One hundred twenty participants were recruited at two schools in a
middle-class suburb in Holland: three classrooms at an elementary school
(9 through 12 years of age, Grades 4 through 6) and three classrooms at a high
school (12 through 15 years of age, Grades 7 through 9). At each grade level,
20 participants were selected randomly from one classroom (10 girls and 10
boys) and were assigned randomly to the subgroups created by the research
design (in each grade, 5 girls and 5 boys received the oral administration first
and the written administration second and vice versa for the other 5 girls and
boys). The average age of the participants was 12.5 years of age ($SD = 1.77$;
range: 9.5 through 15.8 years of age; for boys, $\bar{X} = 12.5$; for girls, $\bar{X} = 12.4$).
Research Design

A split-half within-participants design was used, that is, each participant received half of the 32 SCT-Y items with the standard written instructions and received the other 16 items by the oral mode. All 120 participants thus served as their own control. Participants were not told that they would receive first a written and then an oral mode (or vice versa) to prevent their responses on the first half (oral or written) from being affected by knowing what would come next (written or oral). In addition, the order of the oral and written half of the SCT-Y was counterbalanced: 60 participants received the written mode first and the oral half second and vice versa for the other 60 participants, who received the oral mode first and the written half second (as noted previously, the counterbalancing was done equally within each grade and gender). By this procedure, it would be possible to document change within-participants while controlling for carryover and practice effects. Counterbalancing the order of the presentation mode also created a between-participants design: The 60 respondents receiving the oral version first could be compared with the 60 respondents receiving the written version first and vice versa for the second part.

To check whether the first or the second 16 items of the SCT-Y might be more susceptible to an oral or written procedure, the two halves of the SCT-Y also were counterbalanced: 60 participants received items 1 through 16 of the SCT-Y in the written mode and items 17 through 32 in the oral mode, whereas the other 60 participants received items 17 through 32 in the written mode and items 1 through 16 in the oral mode. The SCT-Y raters were blind to the presentation mode (all responses from both conditions were mixed and put in random order).

Instruments

Level of ego development was measured by means of the SCT-Y (Westenberg, Treffers, et al., 1998). The SCT-Y consists of 32 items, and the respondent is instructed to “complete the following sentences in any way that you wish.” The data were scored by the same procedures used to score the WUSCT (Hy & Loevinger, 1996; Loevinger, 1998): (a) For each respondent, the 32 responses were typed into a spreadsheet; (b) the responses then were sorted by item and were put in random order (i.e., all 120 responses to Item 1 were grouped together and randomized); (c) information on the participant and the conditions was hidden from view to allow for ratings out of context (i.e., blind ratings with respect to age, grade, and written-oral condition); (d) each response was rated by two independent raters using the scoring man-
ual for the SCT-Y (a rating consists of a literal match between the response and a response category in the scoring manual; see Westenberg, Jonckheer, et al., 1998); (e) the relatively few differences between the two raters were resolved by asking a third rater to rate these responses, to resolve the differences by a majority vote; (f) the responses and the ego-level ratings then were resorted to the original protocols, yielding two sets of 16 item ratings for each individual (one set for the oral mode and another for the written mode); and (g) the frequency distribution of the two sets of item ratings was converged into two total scores.

The Total Protocol Rating (TPR; i.e., level of ego development) was computed by converging the cumulative frequency distribution of the item ratings into a discrete ego-level score on the basis of the ogive rules, yielding an oral and a written TPR for each participant. In addition, an Item Sum Score (ISS) was computed by adding the 16 item ratings for both test halves, yielding an oral and a written ISS for each participant. Even though the ISS is not considered to be an adequate index of level of ego development, it was used to maximize the likelihood of finding differences between oral and written administration modes (i.e., a continuous variable is frequently more powerful statistically than is an ordinal variable). Incidentally, the distribution of item ratings was doubled to arrive at 32 item ratings for the written and oral test halves before the ISS and TPR were computed.

The reading/writing skills of the fourth-grade through sixth-grade students were assessed in two ways. First, the teachers were asked for their judgment on whether their pupils might experience significant difficulties in completing the SCT-Y in a written format, rated on a 5-point scale ranging from $1 = \text{no, not at all}$ through $5 = \text{yes, definitely}$. Second, elementary education in the Netherlands includes a standardized system for tracking the language development of the children, consisting of five subscales covering three areas (reading, spelling, and writing). The six available scores (including the teachers’ judgment) were interrelated, as was shown by a Cronbach’s alpha of .89. On the basis of the six scores, the group was divided into three equal-sized groups: a group with relatively low reading/writing skills, an intermediate group, and a group with good reading/writing skills ($n = 20$ in each group). Such reading/writing data were not available for the seventh- through ninth-grade students, but on the basis of the McGammon (1981) findings it was assumed that by the seventh grade all schoolchildren would have developed sufficient writing skills for a written administration of the SCT-Y.

To study the effect of personal preference for either administration procedure, participants were asked two questions (after the SCT-Y had been completed): “Which procedure did you like best (was most fun to do)?” and “Which procedure was easiest for you?” Responses to both questions were
rated on two separate 3-point scales: 1 = the written procedure was liked best or was the easiest, 2 = no procedure was more fun or easier than the other, and 3 = the oral procedure was liked best or was the easiest.

RESULTS

The first issue to be explored was the question of whether oral responses would differ from written responses in terms of the type and the length of the response. Whether the type of the oral responses differed from the written responses was studied in terms of the comparative utility of the scoring manual of the SCT-Y in the oral as compared with the written presentation mode. Could orally elicited responses be classified readily by the current scoring manual developed on the basis of written responses?

In the written mode, an average of 92.8% of the responses (32 items, 60 responses per item) could be classified readily on the basis of the scoring categories provided by the scoring manual for the SCT-Y (range: 75% through 100%). In the oral mode, an even higher average of 95.6% of the responses (32 items, 60 responses per item) could be classified with the scoring manual (range: 80% through 100%). Thus, even though the manual was developed on the basis of written responses, it covered a larger percentage of the oral responses. However, according to a chi-square test that difference was not significant statistically.

Second, the potential difference between oral and written responses was studied also in terms of the length of the response (i.e., the number of words). The average length of a written response (X = 5.31 words) was lower than the average length of an oral response (X = 5.75 words), but that difference was not significant (t = 1.79, ns). Those findings indicate that oral responses are not essentially different from written responses in terms of type or length.

The second set of analyses concerned the reliability of the SCT-Y under both conditions. The first aspect of reliability studied was the interrater agreement between the two independent raters. The average interrater agreement across the 32 items (60 responses per item) for the written condition was 87.4% for perfect agreement (range: 72.4% through 98.3%) and 96.1% if allowing for one-stage disagreements (range: 91.3% through 100%). The interrater agreement for the oral condition was highly similar: 87.7% for perfect agreement (range: 71.7% through 98.3%) and 96.4% if allowing for one-stage disagreements (range: 87.2% through 100%). The Kappa index of interrater agreement ranged from .63 to .98 in the written condition (average Kappa was .78) and ranged from .62 to .98 in the oral condition (average Kappa was .82; all Kappas were significant statistically at the .0001 level). In
summary, the interrater reliability was equally high under both conditions and compared favorably with the level of rater agreement reported in other studies. The relatively few disagreements were resolved by a third rater (see Method section).

The second aspect of reliability concerns the internal consistency of the items under both conditions (Cronbach’s alpha). The internal consistency of the SCT-Y could not be computed across all 32 items because each respondent completed half of the items written and the other half orally. The results showed comparable internal consistency for both presentation modes: Items 1 through 16 displayed an alpha of .81 for the written mode and an alpha of .81 for the oral mode, and Items 17 through 32 displayed an alpha of .75 for the written mode and an alpha of .81 for the oral mode. If corrected for the number of items, these alphas would compare favorably with the alphas reported in the literature for the 36-item WUSCT (see Loevinger, 1998).

The third aspect of reliability concerns the split-half rank-order stability of the ego-level scores. Ego-level scores obtained on the oral half of the SCT-Y correlated significantly with the scores obtained on the written half: .77 for the ISS and .63 for the TPR (p < .0001). The correlations were not different significantly for girls or for boys or for the different age groups. Those correlations indicate that, by and large, individuals maintain their rank-order position across both presentation modes. Indeed, those correlations are hardly lower than the split-half reliability of the WUSCT administered entirely in the written mode. Novy and Francis (1992) reported a correlation of .79 for split-half ISS scores, as compared to a correlation of .77 observed in the present study.

The high utility and reliability of the scoring manual for both presentation modes allowed for the third set of analyses. Those analyses addressed the question of whether different presentation modes would yield differing ego-level scores. The findings indicated that the different modes did not yield differing ego-level scores. First, in both conditions, about half of the participants were at the Self-Protective ego level, about one-quarter were at the Impulsive level, and one-quarter were at the Conformist level. In other words, in the present sample the Self-Protective level was the modal level of ego development, regardless of presentation mode. Only 4 participants (3.3%) scored at the Self-Aware or Conscientious ego levels. Those cases were considered to be outliers and were added to the group of Conformist individuals. Therefore, the TPR scores ranged from 2 = Impulsive level through 4 = Conformist level.

Second, according to a repeated measures ANOVA, the within-participants main effect of presentation mode on ego-level scores was not significant statistically. The TPR and the ISS scores were remarkably similar for both pre-
sentation modes: (a) For the written mode the mean TPR was 3.08 ($SD = 0.76$) and for the oral mode the mean TPR was 3.15 ($SD = 0.67$; $F = 1.81, ns$), and (b) for the written mode the mean ISS was 105.10 ($SD = 16.4$) and for the oral mode the mean ISS was 105.16 ($SD = 15.2$; $F = 0.01, ns$). Across the board, an oral or written mode of presentation did not affect ego-level scores, and the Self-Protective ego level was the modal and the mean level of ego development in the present sample.

The results also were computed separately for the two differing orders of the oral and the written modes (oral-written or written-oral) to investigate possible order effects and study between-participants effects. The results presented in Table 2 indicate neither significant order effects nor significant between-participants effects.

Similarly, no interaction effects were observed for the order of the selected items (i.e., Items 1 through 16 or Items 17 through 32 of the SCT-Y). Therefore, the discussion of the following analyses will be based on the oral/written distinction, without further reference to the order of the items or the order in which the presentation modes were used.

The fourth set of analyses concerned the possible interaction effects of presentation mode by several individual difference variables: age/grade, gender, reading/writing skills, and preference for either presentation mode. As a test of the possible interaction effect of presentation mode by age, participants were grouped into three equal-sized age groups: 9.5 through 11.5, 11.5 through 13.5, and 13.5 through 15.5 years of age ($n = 40$ in each group; see Table 3). The interaction effect of presentation mode by age group was not significant statistically for the TPR but was significant for the ISS ($F = 3.31, p < .05$). However, none of the observed ego-level differences (i.e., ISS ratings) for the oral and written administration modes was significant statistically by separate $t$ tests. The interaction effect of presentation mode by grade level was not significant either, neither for the TPR nor for the ISS. In other words, no age group or grade level benefited significantly from either presentation mode. Naturally, age and grade level were related strongly to level of ego development, regardless of presentation mode. For the written procedure the correlations with age were .65 for the TPS and .67 for the ISS ($p < .0001$). For the oral procedure the correlations with age were .65 for the TPS and .69 for the ISS ($p < .0001$).

The reviewed studies in which the WUSCT was used indicated an interaction effect for gender: Girls and women scored higher on the written mode as compared to the oral mode, whereas no effect of presentation mode was observed for boys and for men (e.g., McGammon, 1981). However, that interaction effect was not tested statistically in those studies. A significant interaction effect of presentation mode by gender was not observed in the
findings of the present study: Neither the girls nor the boys scored significantly higher on the oral or written presentation modes (see Table 3). However, as was to be expected on the basis of previous research with the WUSCT (see Cohn, 1991), girls scored higher than did boys, regardless of presentation mode (see Table 3). Yet that main effect for gender did not reach statistical significance at the .05 level.

An interaction effect of presentation mode by reading/writing skills was not found, either. Information about individual reading/writing skills was available only for the fourth- through sixth-grade students. A repeated measures ANOVA did not yield a significant interaction effect, that is, at each of the three reading/writing levels there was no significant difference between the oral and written halves of the SCT-Y. In addition, reading/writing scores

<table>
<thead>
<tr>
<th>Total Protocol Rating (TPR)</th>
<th>Item Sum Score (ISS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written</td>
<td>Oral</td>
</tr>
<tr>
<td>Group 1</td>
<td>3.12 (0.74)</td>
</tr>
<tr>
<td>Oral</td>
<td>Written</td>
</tr>
<tr>
<td>Group 2</td>
<td>3.10 (0.68)</td>
</tr>
</tbody>
</table>

NOTE: Group 1 = first half of Sentence Completion Test for Children and Youth (SCT-Y) written (n = 60), second half oral; Group 2 = first half of SCT-Y oral, second half written (n = 60).

<table>
<thead>
<tr>
<th>Total Protocol Rating (TPR)</th>
<th>Item Sum Score (ISS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written</td>
<td>Oral</td>
</tr>
<tr>
<td>Boys</td>
<td>3.00 (0.74)</td>
</tr>
<tr>
<td>Girls</td>
<td>3.15 (0.78)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>9.5-11.5</td>
<td>2.45 (0.55)</td>
</tr>
<tr>
<td>11.5-13.5</td>
<td>3.17 (0.70)</td>
</tr>
<tr>
<td>13.5-15.5</td>
<td>3.62 (0.49)</td>
</tr>
<tr>
<td>Total</td>
<td>3.08 (0.76)</td>
</tr>
</tbody>
</table>

Downloaded from http://jea.sagepub.com at Universiteitsbibliotheek Leiden on July 21, 2009
were uncorrelated with the oral/written difference scores ($r = -.07$ for the ISS; $r = .03$ for the TPR). Apparently, individual differences in reading/writing levels do not contribute to a difference between oral and written ego-level scores. In other words, participants with relatively poor reading/writing skills did not benefit from an oral presentation, and respondents with relatively good reading/writing skills did not benefit from a written presentation. However, the reading/writing skills were related positively and significantly to ego-level scores and, surprisingly, were related more strongly to oral ego-level scores (ISS $r = .54$; TPR $r = .57$; $p < .001$) than to written scores (ISS $r = .47$; TPR $r = .42$; $p < .001$). If age and grade level were controlled, the correlation between reading/writing skills and ego-level scores remained significant at the .05 level.

The last set of data analyses concerns the respondent’s personal preference for either presentation mode. Participants were asked which presentation mode they felt was easier (the easy factor) and which one they liked best (the liking factor). Fifty-five participants indicated that the oral procedure was easiest (45.8%) and 53 respondents indicated that they liked the oral procedure best (44.2%). A similar proportion preferred the written procedure: The written procedure was easiest for 52 respondents (43.3%) and was liked best by 50 respondents (41.7%). A minority did not have a preference: 13 for the easy factor (10.8%) and 17 for the liking factor (14.2%).

On the basis of the three age groups used earlier, preference for either procedure was related significantly to age (easy factor: $\chi^2 = 16.49$, $df = 4$, $p < .01$; liking factor: $\chi^2 = 14.37$, $df = 4$, $p < .01$). The oral procedure was preferred by the majority of the youngest age group (9 through 11.5 years of age): 27 members of this group thought it was easier (67.5%) and 26 members liked it best (65%). In contrast, the written procedure was preferred by the majority of the oldest age group (13.5 through 15 years of age): 23 members of this group thought it was easier (57.5%) and 21 members liked it best (52.5%). Interestingly, quite a few of the oldest group preferred the oral procedure: 15 thought it was easier to respond orally (37.5%) and 14 liked it best (35%).

The important question, however, was whether a preference for either administration mode would be related to differing ego-level scores. The results showed no interaction effect of presentation by the easy factor: (a) The oral scores were not higher than the written scores, not even for the participants who felt the oral mode to be easiest; and (b) the written scores were not higher than the oral scores, not even for the participants who felt the written mode to be easiest (same results for ISS and TPR scores). Similarly, no interaction effect was observed for presentation mode by the liking factor, at least not for TPR scores. For ISS scores, however, a significant interaction effect was observed. According to a repeated measures ANOVA, this interaction of
presentation mode by liking was significant at the .001 level ($F = 8.76$). The paired $t$ tests were significant also: (a) For the participants preferring the oral mode, the oral ego level scores ($\bar{X} = 102.51$, $SD = 16.22$) were higher significantly than the written scores ($\bar{X} = 98.23$, $SD = 16.30$; $t = 2.63$, $p = .011$); whereas (b) for the participants preferring the written mode, the written ego level scores ($\bar{X} = 110.52$, $SD = 15.0$) were significantly higher than the oral scores ($\bar{X} = 107.23$, $SD = 14.80$; $t = 2.66$, $p = .011$).

**DISCUSSION**

The findings from the present study indicate that the revised description of the lowest ego levels (Westenberg, Jonckheer, et al., 1998) and the revised scoring manual (i.e., the newly constructed SCT-Y) (Westenberg, Treffers, et al., 1998) are not an artifact of the written procedure used for studying ego development in (young) adolescents. Even though the scoring categories in the manual for the SCT-Y were developed on the basis of written responses, oral responses could be scored readily and reliably with the same scoring manual. In addition, oral responses were not more elaborate than were written responses. Those findings indicate that oral responses are not different essentially from written responses and would not have resulted in a different set of scoring manuals or stage descriptions.

With regard to the potential ego-level differences due to a written or oral presentation, the SCT-Y appears a very robust instrument, at least for this age group (9 through 15 years of age). Based on a split-half within-participants design, it was observed that ego-level scores were not affected systematically by presentation mode. The absence of significant and consistent written/oral differences was observed in all age cohorts, for both genders, at all reading/writing levels, and irrespective of the respondent’s preference. Even the participants with relatively limited reading/writing skills did not obtain significantly lower written, as compared to oral, scores.

The results indicate that the youngest respondents (9 through 12 years of age) did not benefit from an oral presentation and that the oldest respondents (12 through 15 years of age) did not benefit from a written presentation. Yet the majority of the youngest respondents preferred the oral mode, whereas the majority of the oldest respondents preferred the written mode. However, the preference for the oral or written mode was not related consistently to higher ego-level scores on the oral or written half of the SCT-Y, at least not for TPR scores. A significant interaction effect was found, however, for preference by ISS scores. Participants who preferred the oral mode did obtain higher oral ISS scores, whereas participants preferring the written mode did
obtain higher written ISS scores. Apparently, ISS scores are more susceptible to preference than are TPR scores. That finding supports the Loevinger (1998) assumption that TPR scores are less dependent on extraneous influences, such as motivation level and word count, and that ISS ratings need to be used with great caution.

The absence of an interaction effect for gender indicates that the SCT-Y is more robust than the WUSCT, at least for these young adolescents. Studies in which the WUSCT has been used revealed a consistent written/oral difference for females but not for males (i.e., females scored higher on the written half than on the oral half; Hansell et al., 1985; McGammon, 1981; Streich & Swensen, 1985). In contrast, the SCT-Y did not display that bias in favor of the girls for the written mode. On the basis of a meta-analysis, Cohn (1991) reported a main effect for gender: On average, females outscore their male counterparts. A female advantage was observed also in the present study but was not significant statistically. Those findings raised the question of whether the gender differences reported by Cohn (1991) actually might be due to a gender bias in the WUSCT (if presented in the standard, written mode). However, a large, normative study of ego development with Dutch children and adolescents, based on a written presentation of the SCT-Y, displayed a female advantage congruent with the female advantage reported by Cohn (1991; see Westenberg et al., 2000). In large or pooled samples, the female advantage emerges consistently, regardless of the particular instrument (WUSCT or SCT-Y) and regardless of nationality (American or Dutch). The absence of a significant advantage of the girls in the present sample simply might be due to unknown sampling factors. As Cohn (1991) showed through the meta-analysis, not all studies revealed a significant advantage for the girls.

The significant correlations between level of reading/writing skills and ego-level scores (written and oral ISS and TPR) resemble the reported relations with verbal intelligence: Ego level is related moderately and significantly to verbal intelligence (e.g., McGammon, 1981; Newman, Tellegen, & Bouchard, 1998; Westenberg & Block, 1993). However, although ego-level scores were related to verbal intelligence, that relation did not reflect written/oral differences (McGammon, 1981). Similarly, the present findings indicated that the level of reading/writing skills was unrelated to written/oral differences.

It might be argued that the absence of significant mean-level differences for the written and oral conditions might be due to a lack of sufficient power due to splitting the SCT-Y into two, less reliable halves. However, the present findings indicated excellent reliability for the two halves across both presentation modes. Moreover, in other studies in which the same split-half design was used, there has been documentation of significant ego-level differences.
between the two test halves. As noted previously, Drewes and Westenberg (2001) used the two halves of the SCT-Y to study the effect of the instruction to complete the sentence stems “in as adult and mature a manner as you can.” The latter instruction raised the TPR scores by a significant margin, presumably toward the respondent’s optimal ego level. Similarly, Blumentritt, Novy, Gaa, and Liberman (1996) used the two halves of the WUSCT to study the effect of the instruction to complete the sentence stems “the most complex and thought-provoking way that you can.” The latter instruction raised the TPR scores by a statistically significant margin. In other words, the two test halves provide sufficient power for testing potential differences if the instruction is congruent with the ego development construct.

In summary, the findings indicate that the oral and written procedures virtually are interchangeable in terms of the respondents’ responses and ego-level scores, and that they both might be applied as the testing situation requires. Those and other findings with the SCT-Y (e.g., Westenberg et al., 1999) indicate that the Loevinger (1976, 1993) conception of ego development is accessible also in the early adolescent period by means of the sentence completion method. At the same time, however, important modifications are in order (see Westenberg, Jonckheer, et al., 1998). Many young adolescents are at the Impulsive or Self-Protective ego levels but do not fit the Loevinger description of those levels entirely. The majority is not hostile or manipulative predominantly but appears to be vulnerable or self-reliant. Young adolescents at the Conformist ego level do not follow very concrete rules rigidly but appear to adhere to several more abstract rules, such as reciprocity and equality in relationships.

The present findings and the findings from other studies (e.g., Avery & Ryan, 1988; Gfellner, 1986; for an overview see Cohn, 1998) indicate the Self-Protective ego level to be the modal level in the early adolescence period. According to the revised description of ego development, the most important aspect of the transition from the Impulsive to the Self-Protective ego level is the reversal of a basically dependent attitude toward a basically independent attitude (Westenberg, Jonckheer, et al., 1998). The Impulsive ego level is characterized by vulnerability and dependency on parents or other caretakers for impulse control and protection, and care generally. In contrast, the Self-Protective level is characterized by a sense of invulnerability and an emphasis on self-control and self-reliance. The Impulsive to Self-Protective transition fits the repeated observation that early adolescence is marked by a normative increase of independency (i.e., emotional autonomy and individuation; see Steinberg, 1999). Development from the Impulsive to the Self-Protective ego level provides an additional paradigm for studying the
course and dynamics of increasing independency in the early adolescence period.

In addition, the ego development paradigm also contributes to understanding of individual differences within the early adolescence period. Individual differences might be due, in part, to differential maturity, but not in a simple, linear manner. For example, independency is not just low at low ego levels and high at high ego levels, or vice versa, but is expected to display a non-linear pattern. On the basis of the revised conception of ego development, independency is expected to peak at the modal, Self-Protective ego level. In contrast, relatively immature young adolescents at the Impulsive ego level and the relatively precocious individuals at the Conformist ego level are expected to be less independent, albeit in differing ways and for differing reasons. The dependency of the Impulsive person is fueled by a basic sense of physical vulnerability, whereas the dependency of the Conformist individual is fueled by the need for belonging.

The trait approach to personality inevitably would confuse both types of dependency, mixing them up at the low end of the independency scale (or at the high end of a dependency scale). The developmental psychologist would not overlook the qualitative distinction between both kinds of dependency but might be inclined to overlook substantial differences in developmental maturity. It is the merit of the Loevinger (1976, 1993) conception and measure of ego development that it at once captures a developmental sequence and a dimension of individual differences in any age cohort (see Westenberg, Blasi, et al., 1998).

NOTE

1. The Washington University Sentence Completion Test (WUSCT) does include a separate protocol for children and adolescents (Form 2-77) (see Loevinger, 1998), but some of the items do not have a formally validated and revised scoring manual and none of the items or scoring manuals were validated specifically for use with children and adolescents. For a comparison between the items of Form 2-77 and those of the Sentence Completion Test for Children and Youth (SCT-Y), see Westenberg, Treffers, and Drewes (1998).

REFERENCES


Requests for reprints should be addressed to P. Michiel Westenberg, Ph.D., Leiden University, Faculty of Social Sciences, Developmental and Educational Psychology, Wassenaarseweg 52, Postbus 9555, 2300 RB Leiden, The Netherlands; e-mail: westenberg@fsw.leidenuniv.nl