5. RESULTS

After a brief description of statistical procedures and data screening, results are presented in three sections. The first section (Section 5.2) presents the effects of instruction and adaptive training on task performance in the two cognitive tasks (N-back and MOL) and the balance task. Age groups initially differed in cognitive task performance, but the adaptive phase successfully reduced age differences when performance was measured in percent correct. Section 5.3 reports that single-task baselines in all three tasks were stable over the course of the dual-task phase. Section 5.4 presents age differences in task performances under the single- and dual-task conditions and introduces the calculation of proportional dual-task costs.

Empirical support for the first hypothesis – namely that children show a more pronounced performance trade-off in favor of the sensorimotor domain than adults – is presented in Section 5.4.2. However, this trade-off pattern was not more pronounced in children if the difficulty of the balance task was increased, as shown in Section 5.4.3. The final section (Section 5.5) presents performance under differential-emphasis conditions, demonstrating that children's trade-off pattern in favor of the sensorimotor task prevailed even when they were instructed to focus more strongly on the cognitive task.

5.1 Statistical Procedures and Data Screening

Mixed-design analyses of variance (ANOVAs) with age group (2) as between-subjects factor were used. Planned contrasts were specified for the group factor comparing (a) young adults to the mean of the two children's groups, and (b) 9- to 11-year old children. The nature of the within-subjects factors depended on the specific analysis (e.g. "trial" for the analyses of single-task baselines, or "task domain" for the analyses of dual-task costs). Linear and quadratic trends were often specified as contrasts for the within-subjects factors. To assure that Type I error and power could be interpreted, several assumptions have to be met in these analyses. The first assumption refers to the independence of observations, which means that one participant's scores should not be influenced by another participant's scores. This assumption was met by study design. Second, data should be sampled randomly from the population and measured at interval level scales. For the present study, participants had been selected on the basis of their cognitive performances in psychometric tests during a screening protocol. Therefore, results can only be generalized to comparable populations. Third, interval level of measurement scales was assumed for the variables under investigation. The fourth assumption is that of multivariate normality, indicating that the distribution of the dependent variables should show multivariate normality within groups. Following the recommendation by Field (2000), the assumption of univariate normality was assessed for each dependent variable separately by Bartlett's Box test. Because sample sizes in this study were rather small, violations of this assumption did occur. However, the ANOVA is known to be rather robust against violations of the normality assumption (as discussed in Bortz, 1999 or Stevens, 2002), and since the sample sizes were comparable between groups, the data of the current study have not been transformed to normalize data distributions.

Finally, another assumption for repeated measures ANOVAs is that of sphericity. That assumption concerns the relationship between scores at all levels of the within-subjects variable. Sphericity is a less restrictive condition of compound symmetry, and data are compound symmetric if the variances of the measures at each level of the repeated factor are equal and the covariances between the measures at each level of the repeated factor are also equal. In other words, sphericity refers to the equality of variances of the differences between treatment levels (see also Field, 2000, and Grimm & Yarnold, 2000). The extent to which the covariance matrix deviates from sphericity is reflected in the parameter ε (Epsilon). For $\varepsilon = 1.0$, sphericity is met, while the value for the worst possible violation of shericity is $\varepsilon = 1/(k - 1)$, where k is the number of treatments. If the assumption of sphericity is not met, adjusted F-tests can be used, for which the numerator and denominator degrees of freedom are reduced. The rather conservative Greenhouse-Geisser (1959) correction has been used in cases where the data departed significantly from sphericity. Because this correction did not lead to changes in the significance of the effects under investigation, results were reported without the correction. The value for η^2 (partial Eta squared) is presented for each effect, which describes the proportion of total variability attributable to the factor.

Cronbach's Alpha reliability coefficients were calculated for each variable under investigation. Tables E.1 to E.3 with the coefficients are presented in Appendix E along with a discussion of the findings concerning reliability. Overall, reliabilities were satisfactory or high, with few exceptions. Reliability coefficients were higher for the raw scores as compared to the dual-task costs.

For the analysis of the dual-task performance, data were analyzed at the level of raw scores as well as at the level of dual-task costs. For the raw scores, dependent variables were the percent-correct scores for the two cognitive tasks (MOL and N-back), and square-root transformed COP areas for the balance task. The dual-task costs were expressed in a proportional metric, by presenting the performance-reductions under dual-task conditions in percent of single-task performance (see Section 5.3.1 for details). Significance levels in all analyses were fixed at $\alpha \leq .05$.

For the data set of the current study, a combination of different procedures was used to detect *outliers*. In SPSS, box plots were used to graph individual trials of all the participants belonging to a specific age group, including all the trials conducted under comparable conditions. For example, all adaptive, single- and dual-task balance trials on the stable platform with a certain trial length (e.g., 35 seconds) were included in one box plot for each age group. Box plots present the median, quartiles, and extreme values of the distribution of the variable. The box represents the interquartile range containing 50 % of the values. Extreme values are defined as cases with values more than 3 box lengths from the upper or lower edge of the box. Additionally, the procedure described by Stevens (2002) for the detection of outliers was used, in which variables are z-transformed, and cases with z-values > 3 are considered as outliers. These are cases more than 3 standard deviations above or below the mean of that group. The results of these two outlier detection procedures were combined, such that only cases detected as extreme values or outliers in both procedures (SPSS box plots and z-transformation) were considered as outliers in the current study. One male young adult produced outliers and extreme values in all of his balance trials on the moving platform. He was replaced by another young adult. For the remaining participants, outliers were only found in the balance task, and not in the MOL or N-back tasks.

Table 8. *Balance Outliers in the Different Conditions of the Study*

	9-year olds	11-year olds	young adults
35 Second Trials on Stable Platform	ID 2001, 1 trial DT		
60 Second Trials on Stable Platform		ID 4006, 2 trials ST	
35 Second Trials on Moving Platform		ID 4003, 1 trial adaptive	
60 Second Trials on Moving Platform			
Dual-Task Costs	ID 2003, stable with numbers		ID 9009, stable with numbers

Note. Outliers are defined as showing extreme values in the SPSS box plots procedure and to additionally be more than 3 standard deviations (SDs) above the mean of the age group (z-value >3). “ST” = single-task; “DT” = dual-task.

Table 8 gives an overview of these cases. For the analysis of specific conditions in which outliers occurred, their values were replaced by the second-highest value from the respective age group. This way of treating outliers is preferable to exclusion (Stevens, 2002). Analyses were also conducted with outliers included, and these results are reported in footnotes in case there were discrepancies to the data set with replacement.

5.2 Effects of Instruction and Adaptive Training on Task Performance

The central question addressed in this section is whether instruction and adaptive training successfully equalized single-task performances in the cognitive tasks across age groups. The section presents the results for each task separately: for MOL, for N-back, and for balance. For the balance task, performance improvement with practice and the influence of difficulty manipulation on performance are investigated.

5.2.1 Instruction and Adaptive Training in the MOL Task

To investigate the effect of task instruction in the MOL task, performance on the pretest was compared to performance on the second block of MOL training, in which the same task parameters were used.

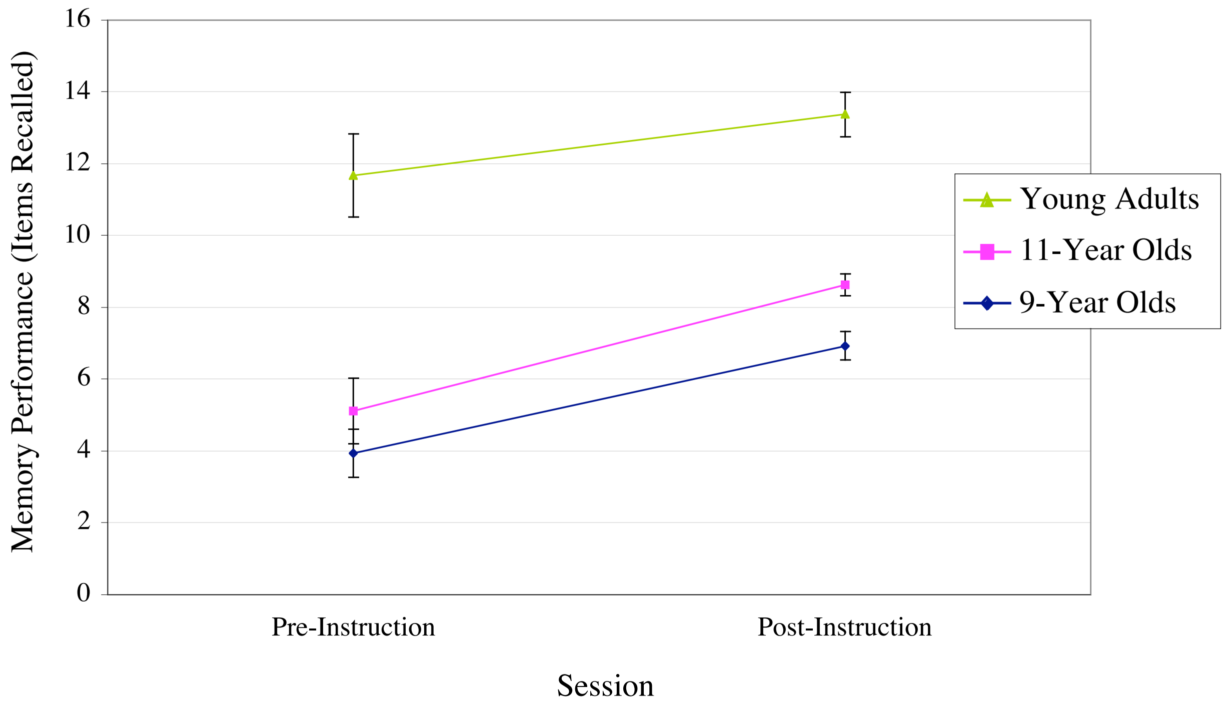

Figure 9. All Age Groups Improve Performance After Instruction in the Method-of-Loci Task

Note. Error bars depict standard errors of the mean.

Figure 9 depicts the pattern for the three age groups. A mixed-design ANOVA was conducted, with age group (2) as between-subjects factor and time of test (2) as within-subjects factor, and with the number of items remembered correctly as the dependent variable. Concerning age differences in memory performance, children differed significantly from young adults, $F(1,24) = 91.16$, $MSE = 5.33$, $p < .001$, $\eta^2 = .792$, while the difference between 9- and 11-year old children did not reach significance, $F(1,24) = 3.48$, $MSE = 5.33$, $p = .075$, $\eta^2 = .126$. The main effect of instruction was significant, $F(1,24) = 22.34$, $MSE = 4.52$, $p < .001$, $\eta^2 = .482$, and it did not interact with the age contrasts, suggesting that the three age groups benefited equally from instruction in the method of loci strategy.[13]

During the adaptive phase of the study, task difficulty was systematically adjusted by changing list length and ISI until each participant remembered approximately 80 % of the stimuli presented. Figure 10 shows the number of words recalled for the three age groups over the course of the adaptive phase.

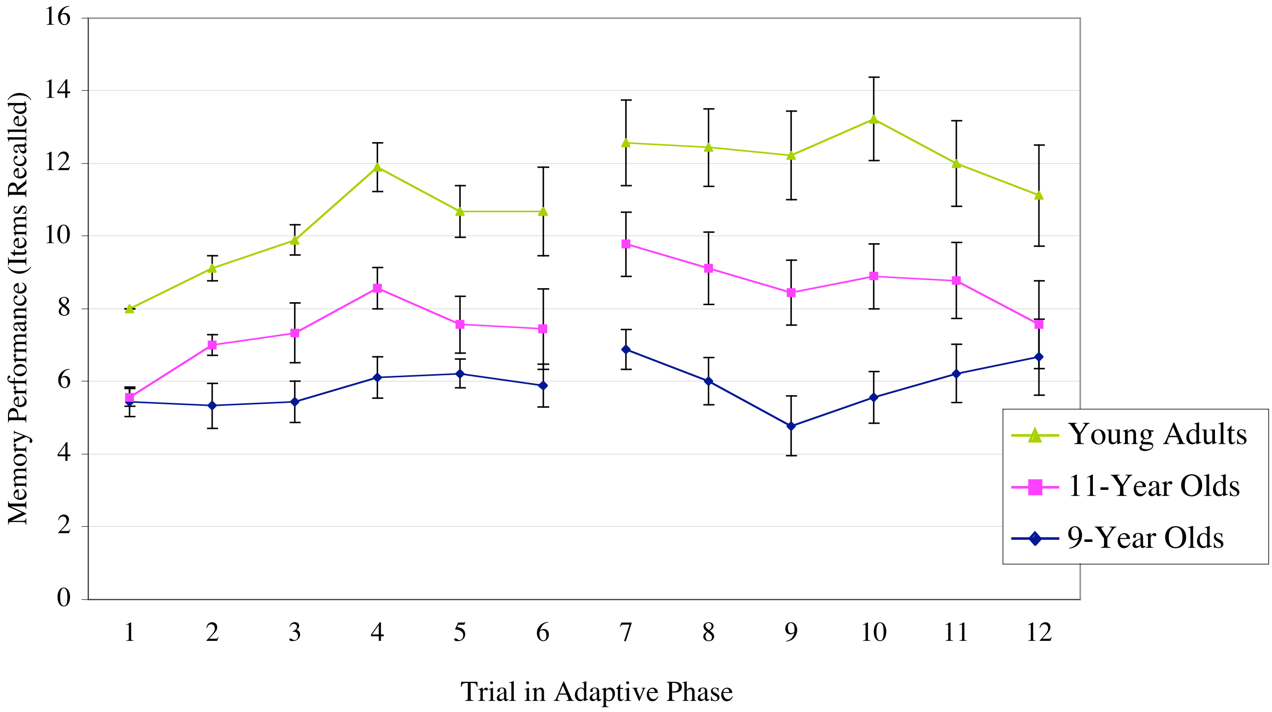

Figure 10. Performance Raw Scores for the Method-of-Loci Task in the Adaptive Phase

Note. Error bars depict standard errors of the mean. The adaptive phase consisted of two sessions, with six trials each.

[13] With fixed list lengths, there were upper limits to how many items could be remembered on each trial. Potential ceiling effects would therefore limit the interpretability of the findings. However, screening the data for such instances revealed that this was not a problem with the current design.

A mixed-design ANOVA with age group (2) as the between-subjects factor and trial within the adaptive phase (12) as within-subjects factor was conducted. Linear and quadratic trends were specified for the trail factor. The age contrasts showed that children remembered significantly fewer words than young adults, $F(1,24) = 30.83$, $MSE = 41.36$, $p < .001$, $\eta^2 = .562$, and the difference between 9- and 11-year olds also reached significance, $F(1,24) = 5.87$, $MSE = 41.36$, $p < .05$, $\eta^2 = .197$, with 11-year olds remembering more words than 9-year olds. For the within-subjects effect of trial, the linear trend reached significance, $F(1,24) = 12.39$, $MSE = 11.34$, $p < .01$, $\eta^2 = .340$, and none of the age contrasts interacted with this trend. This indicates that performance improved further in all age groups during the adaptive phase. In addition, the quadratic trend reached significance, $F(1,24) = 18.56$, $MSE = 4.15$, $p < .01$, $\eta^2 = .436$, and this trend did not interact with the age contrast comparing children to adults, $F(1,24) = 3.04$, $MSE = 4.15$, $p = .094$, $\eta^2 = .112$, but it interacted significantly with the age contrast comparing 9- to 11-year olds, $F(1,24) = 4.48$, $MSE = 4.15$, $p < .05$, $\eta^2 = .157$.

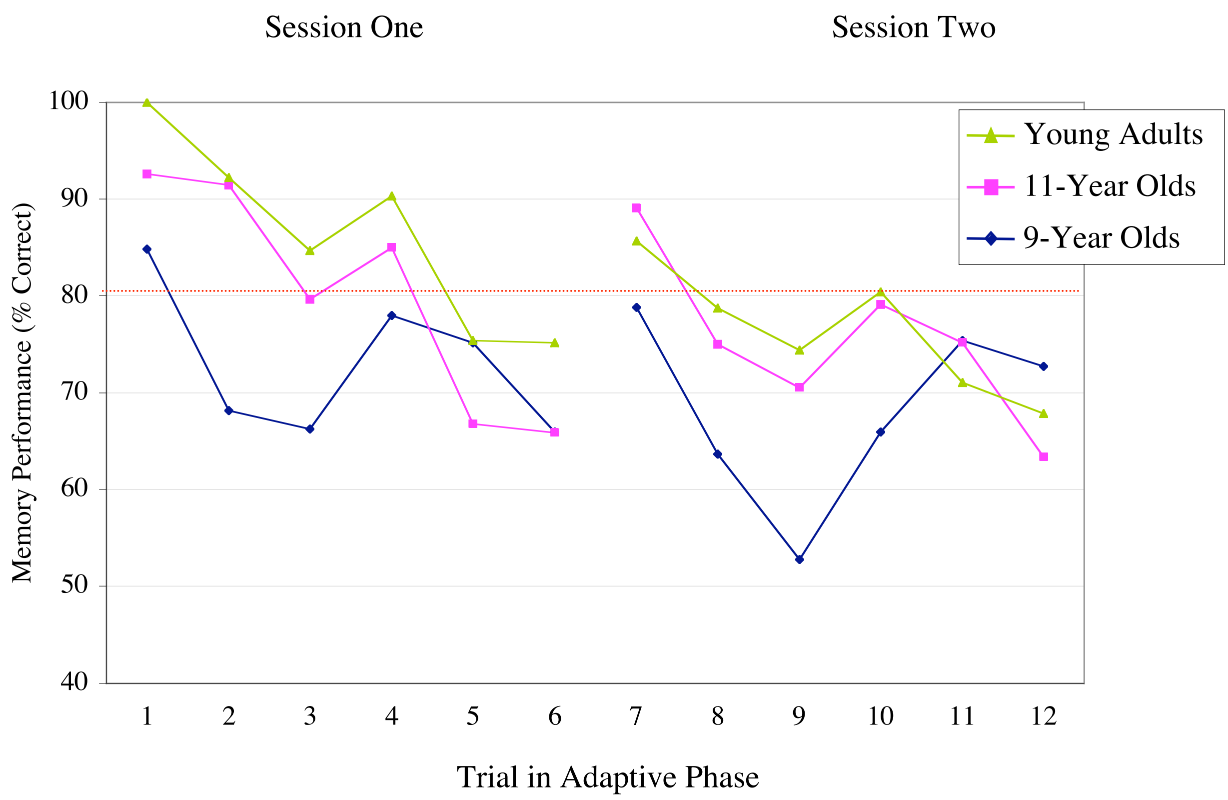

Figure 11. Percent Correct Performance Scores for the Method-of-Loci Task in the Adaptive Phase

Note. The dotted red line represents the 80 % correct performance level.

Figure 11 depicts the data pattern of the adaptive phase when performance scores are expressed in percent correct. This figure indicates that all three age groups approach the 80 % correct level during successive trials. At the end of the adaptive phase, in trial 12, performance (% correct) did not differ significantly between the three groups, as indicated by an ANOVA with age group (3) as the between-subjects factor, $F(2,24) = .479$, $p = .625$. To investigate this result in some more detail, independent samples t-tests were conducted. The t-test comparing children to young adults did not reveal any significant differences between these groups, $t(25)= .026$, $p = .980$, and the differences between 9- and 11-year old children were also not significant, $t(16)= -.47$, $p = .644$. Thus, the efforts to equalize relative performance levels in the three age groups were successful. Appendix F presents additional statistical analysis of the performance changes over the course of the adaptive phase.

5.2.2 Instruction and Adaptive Training in the N-back Task

The N-back task was introduced to participants in the first session, and they practiced it in Sessions 2 and 3 with fixed ISIs and fixed trial lengths. In Sessions 4 and 5, the adaptive phase, task difficulty was manipulated by presenting more items with shorter ISIs (just like in the MOL task). Each adaptive session consisted of two blocks with three trials each. Figure B.1 in Appendix B depicts the performance scores of the three age groups in these two adaptive sessions. Scores for the three trials of each block were averaged. The performance of young adults and children cannot be compared, because young adults worked on the more difficult N-back 4 (instead of the N-back 2) version of the task. Therefore, the following analysis focuses on the performances of 9- and 11-year old children.

A mixed-design ANOVA on the performance scores (correct responses) of the adaptive phase with age group (2) as the between-subjects factor and adaptive block (1 to 4) as the within-subjects factor did not reveal significant differences between the performances of 9- and 11-year olds, $F(1,16) = 2.11$, $MSE = 76.42$, $p = .166$, $\eta^2 = .116$. The linear trend for the effect of adaptive block reached significance, $F(1,16) = 4.72$, $MSE = 12.61$, $p < .05$, $\eta^2 = .228$, and this effect did not interact with the age contrast, suggesting that the children could further improve their N-back performance over the course of the adaptive phase.

The N-back scores expressed in percent correct are depicted in Figure 12. An ANOVA on the percent correct score of the last adaptive block of Session 5 was conducted, with age group (3) as the between-subjects factor. Differences between the age groups were not significant, $F(2,24) = 1.81$, $p = .184$, with children not differing from young adults ($p = .070$) and 9-year olds not differing from 11-year olds ($p = .839$). This shows that performance expressed in percent correct did not differ between the three age groups at the end of the adaptive phase. Appendix G presents the task parameter settings that were chosen for each age group for the dual-task phase of the study.

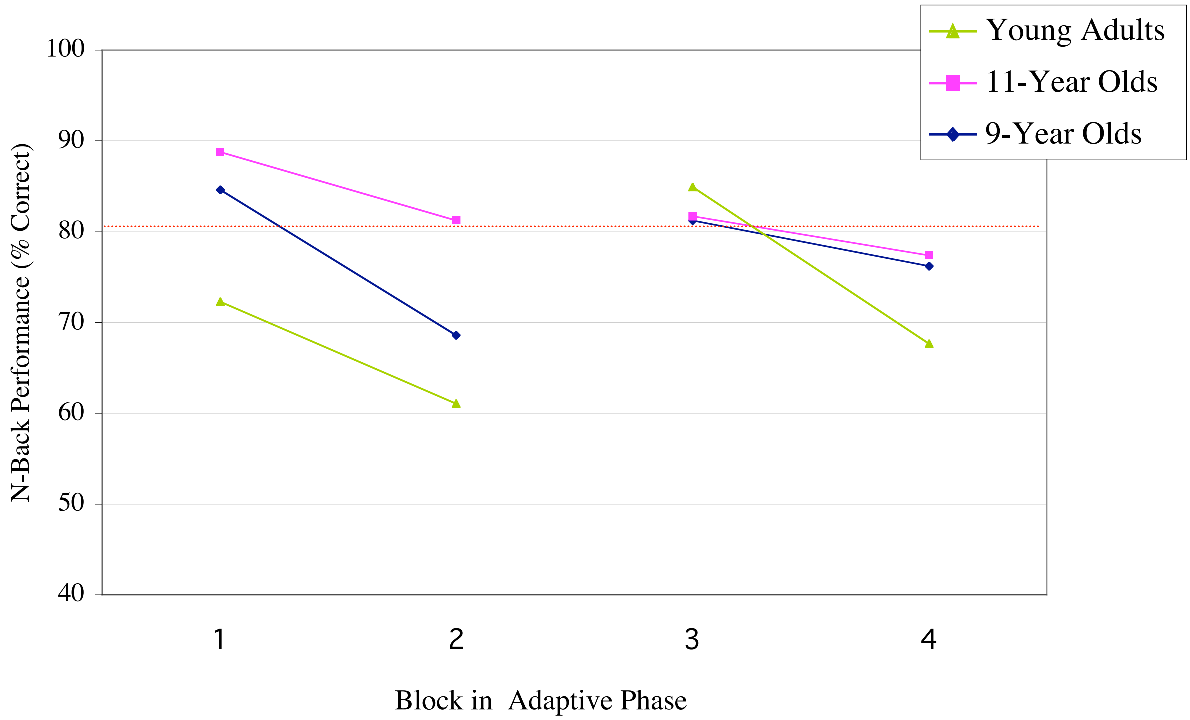

Figure 12. Percent Correct Performance Scores for the N-back Task in the Adaptive Phase

Note. The dotted red line represents the 80 % correct performance level.

5.2.3 Balance Task: Training Effects and Difficulty Manipulation

For the balance task, a square-root transformation was conducted on the COP area measures for each individual trial. Several square-root transformed trials of the same condition were averaged, and the mean was then squared again to obtain measures in the same numerical range as individual trials.

Table 9. *Balance Center-of-Pressure Area Scores for Sessions 1 and 2*

	9-year olds	11-year olds	young adults
Session 1, Block 1			
M	462.22	253.37	193.36
SD	146.06	63.50	43.55
Session 1, Block 2			
M	536.78	310.42	217.66
SD	284.03	157.05	73.43
Session 2, Block 1			
M	421.69	293.36	151.76
SD	168.42	159.26	63.67

Note. COP area scores for each condition have been square-root transformed, averaged, and than squared again.

Means (Ms) and standard deviations (SDs) for average COP areas for the first block of Session 1 (three trials with online feedback), for the second block of Session 1 (three trials with offline feedback), and for the four offline trials of Session 2 are presented in Table 9.

For the data of Session 1, a mixed-design ANOVA with age group (2) as the between-subjects factor and online vs. offline feedback (2) as the within-subjects factor demonstrated that children showed significantly larger COP areas than young adults, $F(1,24) = 12.47$, $MSE = 33000.55$, $p < .01$, $\eta^2 = .342$, and 9-year olds showed larger COP areas than 11-year olds, $F(1,24) = 12.91$, $MSE = 33000.55$, $p < .01$, $\eta^2 = .350$. The comparison of the online feedback block to the offline feedback block of Session 1 deals with the question whether online feedback initially helped participants to stabilize themselves on the ankle-disc board. Table 9 shows that the mean COP areas are smaller for the online feedback condition. However, the ANOVA did not reveal significant differences between the two conditions, $F(1,24) = 2.81$, $MSE = 12999.07$, $p = .107$, $\eta^2 = .105$, and there were no interactions of this effect with the age contrasts.

To investigate whether participants improved their performance with practice during this initial phase of the study, the offline feedback trials of the first session were compared to the offline feedback trials of the second session. A mixed-design ANOVA with practice-block (Session 1 vs. 2) as the within-subjects factor and age group (2) as the between-subjects factor was conducted, and it revealed a nonsignificant trend for the practice effect, $F(1,24) = 3.14$, $MSE = 18751.30$, $p = .089$, $\eta^2 = .116$, which did not interact with the age contrasts. On the level of overall age differences, children and young adults differed significantly in their COP areas, $F(1,24) = 13.59$, $MSE = 37420.76$, $p < .01$, $\eta^2 = .362$, and 9-year olds differed significantly from 11-year olds, $F(1,24) = 7.56$, $MSE = 37420.76$, $p < .05$, $\eta^2 = .240$.

In the third session, moving platform conditions were introduced. Although two different movement conditions were used in each age group, this report will focus on the movement condition that was identical across age groups, namely the 3 degree movement.[14] Figure 13 depicts the average performance of the three age groups on the stable and on the moving platform.

[14] An outlier was detected in the 11-year olds in the moving platform condition (average COP area more than 4 SDs above the mean of that age group). An investigation of individual trials revealed that this was not due to one single extreme trial, but that several trials were considerably higher than that of other individuals of that group. The outlier was therefore adjusted, by replacing it's value by the second largest value obtained in the age distribution. However, when the analysis was conducted without adjusting for the outlier, the same pattern of findings emerged.

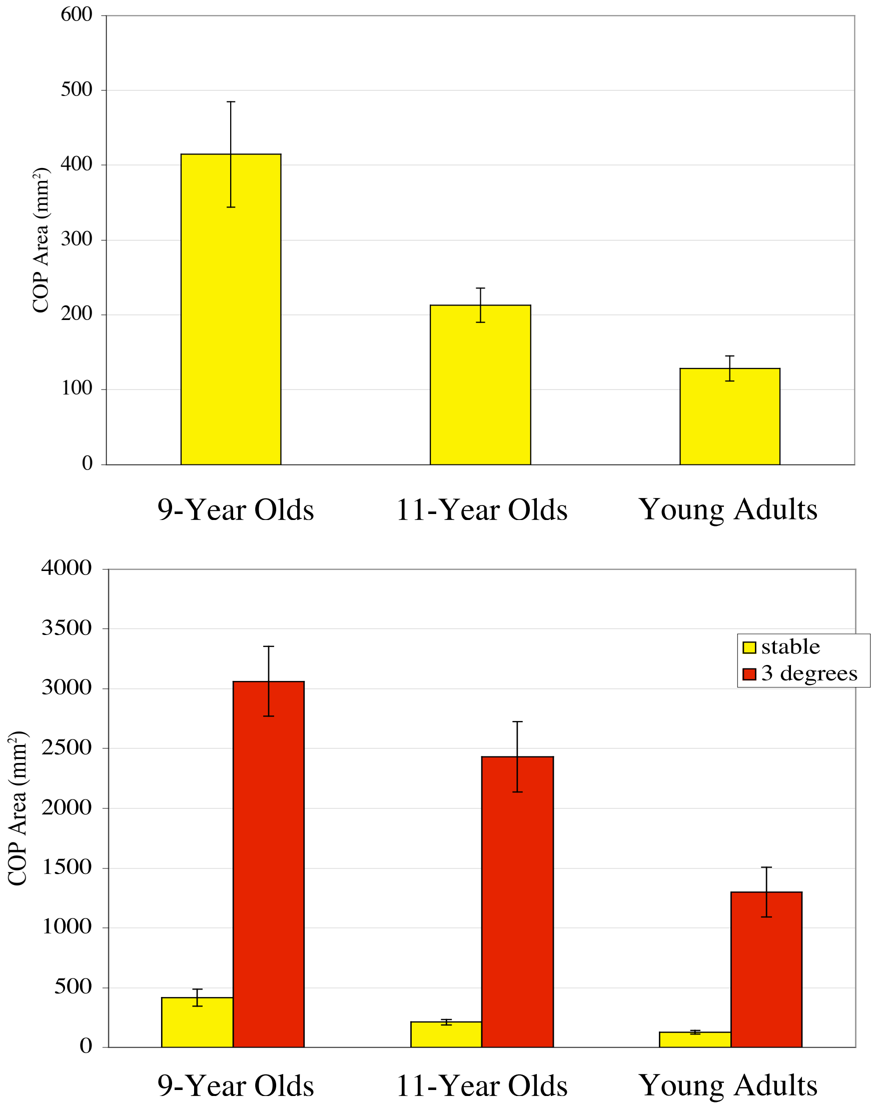

Figure 13. Children Sway More Than Adults, and Sway Increases on the Moving Platform

Note. Error bars depict standard errors of the mean.

A mixed-design ANOVA with age group (2) as the between-subjects factor and balance difficulty (2; stable versus moving platform) was conducted, and it revealed significant differences between the COP areas of children and young adults, $F(1,24) = 20.46$, $MSE = 390331.58$, $p < .01$, $\eta^2 = .460$, while the difference in balance performance between 9-year olds and 11-year olds failed to reach significance by a slight margin, $F(1,24) = 3.99$, $MSE = 390331.58$, $p = .057$, $\eta^2 = .142$. The within-subjects effect of balance difficulty was significant, $F(1,24) = 202.47$, $MSE = 269696.37$, $p < .01$, $\eta^2 = .894$, with COP areas on the moving platform being considerably larger than on the stable platform, and this effect interacted significantly with the age contrast comparing children to young adults, $F(1,24) = 17.69$, $MSE = 269696.37$, $p < .01$, $\eta^2 = .424$, while the interaction with the age contrast comparing 9- to 11-year olds did not reach significance, $F(1,24) = 1.54$, $MSE = 269696.37$, $p = .226$, $\eta^2 = .060$. This indicates that the difference between moving and stable platform was larger for the children than for the young adults.

In Sessions 4 and 5, the balance task was performed either with the secondary task of reading aloud the numbers on the computer screen or with the secondary task of listening to human and animal voices in an auditory shadowing task. Balance was assessed on the stable and on the moving platform with four trials in each task condition. Table 10 presents the means and standard deviations for the balance performances of Session 4 and 5.

Table 10. *Center-of-Pressure Areas for Different Secondary Tasks in Sessions 4 and 5*

		Sample		
Secondary Task	Balance Condition	9-year olds	11-year olds	young adults
Reading Numbers	stable			
	M	307.76	196.62	110.87
	SD	85.65	59.47	18.92
	moving			
	M	1909.72	1130.73	903.53
	SD	413.30	337.86	223.96
Detecting Animal Voices	stable			
	M	607.34	407.35	183.61
	SD	303.42	194.00	66.29
	moving			
	M	2992.70	1962.37	1171.21
	SD	1174.06	955.32	338.74

Note. COP area scores for each condition have been square-root transformed, averaged, and then squared again.

A mixed-design ANOVA with baseline task condition (2; reading numbers versus detecting animal voices) as within-subjects factor and age group (2) as between-subjects factor was conducted for the stable platform condition. As before, significant differences were found between the balance performance of children and young adults, $F(1,24) = 20.51$, $MSE = 31632.61$, $p < .01$, $\eta^2 = .461$, and between 9- and 11-year olds, $F(1,24) = 6.89$, $MSE = 31632.61$, $p < .05$, $\eta^2 = .223$, with older age groups showing superior balance performance than younger age groups. The within-subjects effect of secondary task also reached significance, $F(1,24) = 30.34$, $MSE = 16808.90$, $p < .01$, $\eta^2 = .558$, such that COP areas were larger for the secondary task of detecting animal voices than for the secondary task of reading numbers. That effect interacted with the age contrast comparing children to young adults,

$F(1,24) = 5.94$, $MSE = 16808.90$, $p < .05$, $\eta^2 = .198$, but not with the age contrast comparing 9- to 11-year olds, $F(1,24) = 1.06$, $MSE = 16808.90$, $p = .314$, $\eta^2 = .042$.

The same kind of analysis has been conducted for the data on the moving platform, and it yielded comparable results: Significant differences were found for the COP areas of children versus young adults, $F(1,24) = 16.24$, $MSE = 683027.06$, $p < .01$, $\eta^2 = .404$, and for the COP areas of 9- versus 11-year olds, $F(1,24) = 10.78$, $MSE = 683027.06$, $p < .01$, $\eta^2 = .310$. In addition, the within-subjects effect of secondary task was significant, $F(1,24) = 30.98$, $MSE = 230613.27$, $p < .01$, $\eta^2 = .563$, and it interacted with the age contrast comparing children to young adults, $F(1,24) = 6.19$, $MSE = 230613.27$, $p < .05$, $\eta^2 = .025$, but not with the age contrast comparing 9- to 11-year olds ($p = .440$). When interpreting these differences between trials in which different secondary tasks have been used, it should be kept in mind that it is not only the secondary tasks that differ, but also trial lengths (35 seconds for trials with reading numbers, and 60 seconds for trials with animal voices). Therefore, larger COP areas in the latter condition can be due to the influence of secondary task and the influence of trial length, because longer trials generally lead to larger COP areas. To disentangle these two influences, trial lengths would have to be made comparable by only analyzing the first 35 seconds of the 60-second trials.

Instead of analyzing the average of each balance condition, it is also possible to analyze each trial separately, in order to find out whether participants further improved their balance performance over the course of Sessions 4 and 5. However, there was no indication that participants systematically improved their performance. Detailed information on the results of each balance condition is provided in Appendix H.

5.3 Stability of Single-Task Performance Baselines

This section investigates whether the single-task baselines of the three tasks remained stable over the course of the dual-task phase, or whether there was a systematic change in task performance with time, for example due to increasing levels of practice. For the MOL and the balance task, performances were analyzed on the level of each individual trial, while the analysis of N-back data focused on each block of trials (averaging over two or three trials). Reliabilities of the single-task performances of the three tasks were consistently high, except in the 9-year olds when balancing on the stable platform while listening to animal voices. The coefficients are presented in Table 11.

Table 11. *Reliability Coefficients for Single-Task Baseline Scores*

Task Condition	Number of Trials	Sample			
		9-year olds	11-year olds	young adults	total sample
MOL Single-Task	8	.96**	.96**	.96**	.98**
N-Back Single-Task	18	.82**	.96**	.95**	.98**
Balance Single-Task, Stable with Numbers	8	.83**	.93**	.84**	.96**
Balance Single-Task, Moving with Numbers	4	.85**	.82**	.93**	.92**
Balance Single-Task, Stable with Animals	3	.12	.92**	.90**	.84**
Balance Single-Task, Moving with Animals	3	.81**	.59	.89**	.87**

Note. *$p < .05$, **$p < .01$

5.3.1 Stability of MOL Single-Task Performance

Task difficulty of the MOL task was adjusted individually to 80 % correct in the adaptive phase of the study. Did this level remain stable over the last four sessions, or did participants further improve their performance? Figure 14 depicts the MOL percent correct performance over the course of the dual-task phase. It is obvious that there are marked differences between the first and the second single-task trial in each session (trials 1 vs. 2, 3 vs. 4, 5 vs. 6, and 7 vs. 8). These differences were significant in a mixed-design ANOVA with the within-subjects factor of trial (2), contrasting the average of the first trial of the four sessions versus the average of the second trial of the four sessions, $F(1,24) = 29.21$, $MSE = 44.83$, $p < .001$, $\eta^2 = .549$, and this factor did not interact with the age contrasts. Furthermore, the age groups did not differ in their average percent correct performance.

If one reconsiders the procedure for data collection in the dual-task sessions, namely that the single-task assessment took place at the beginning and the end of each session, it can be speculated that the decrease of MOL performance in the last trial of each session was caused either by increasing tiredness of the participants, or by increases in MOL-specific proactive interference. Proactive interference refers to the phenomenon that images from previous lists might interfere with the creation of new mental images in latter lists, leading to intrusions in recall performance (Kliegl & Lindenberger, 1993).

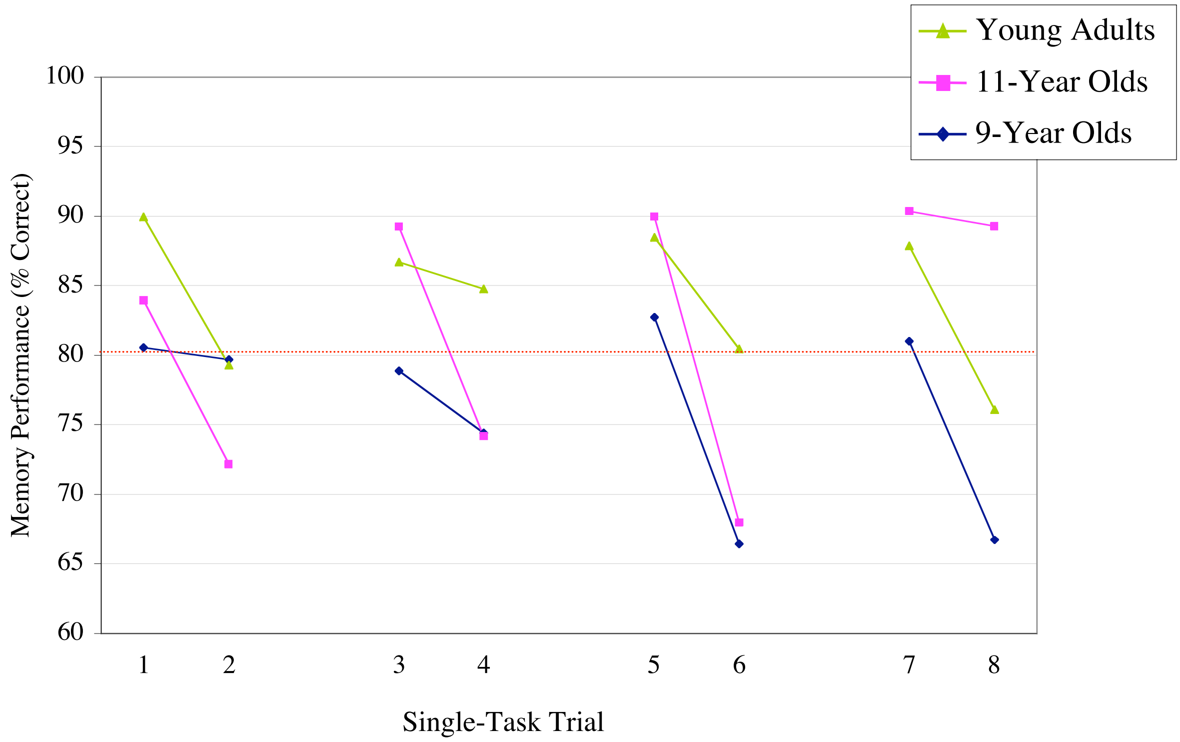

Figure 14. Changes in Method-of-Loci Performance Over the Course of Single-Task Assessment

Note. MOL performance is measured in percent correct. There are two MOL single-task trials in each session, one at the beginning and one at the end of the session. The dotted red line represents the 80 % correct performance level.

To disentangle the influence of tiredness and proactive interference, the results of the two-choice reaction-time task were investigated. Two blocks of this task have been administered at the beginning and end of Sessions 2 to 8. Data of the second session were not included in the analysis, because they were considered practice trials. Figure 15 depicts the reaction times (RTs) of the first and second block of each session for Sessions 3 to 8. The first block has always been assessed right at the beginning of each session, and the second block at the very end of each session. Analysis focused on the RT data, since participants committed almost no omission errors and very few intrusion errors. In a mixed-design ANOVA with block (2; beginning vs. end of session) as within-subjects factor and age group (2) as between-subjects factor, the effect of block did not reach significance, $F(1,24) = .05$, $MSE = 130.66$, $p = .828$, $\eta^2 = .002$. This effect did not interact with the age contrast comparing children to adults, but it did interact with the age contrast comparing 9- to 11-year olds, $F(1,24) = 6.30$, $MSE = 130.66$, $p < .05$, $\eta^2 = .208$. This was due to the 11-year old children sometimes reacting faster in the second block of the session than in the first block. In addition, the age contrasts showed that on the overall level, young adults reacted significantly faster than children, $F(1,24) = 53.85$, $MSE = 2847.39$, $p < .001$, $\eta^2 = .692$, and 11-year olds reacted faster than 9-year olds, $F(1,24) = 6.30$, $MSE = 2847.39$, $p < .05$, $\eta^2 = .202$. These findings suggest no effect of increasing tiredness in the reaction-time data. The performance decrease in the MOL task from the first to the last trial of each session is therefore more likely to be caused by MOL-specific factors like proactive interference.

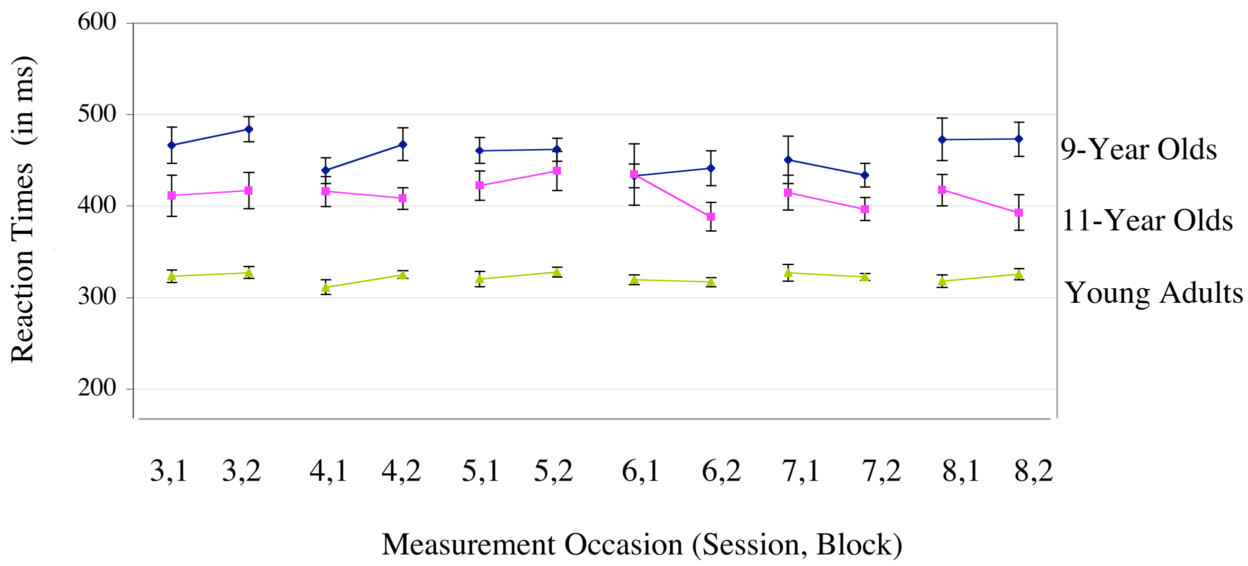

Figure 15. Reaction-Time Performance Does Not Change Systematically Over Sessions

Note. The X-axis depicts the mean reaction time for the two trials of each block, for Sessions 3 to 8. Block 1 comprises the two trials at the beginning of each session, and block 2 the two trials at the end of the session. Error bars depict standard errors of the mean.

5.3.2 Stability of N-back Single-Task Performance

The following section investigates the stability of the N-back baseline performance. Did the N-back performance remain stable at about 80 % correct throughout the single-task trials in Sessions 6 to 9 of the study, or did the performance further improve with practice? As described above, N-back was assessed twice in each dual-task session, before and after the dual-task assessment, with two trials on each occasion. These two trials were averaged, to obtain the mean performance for each single-task block. In the ninth session, each single-task block consisted of 3 trials.

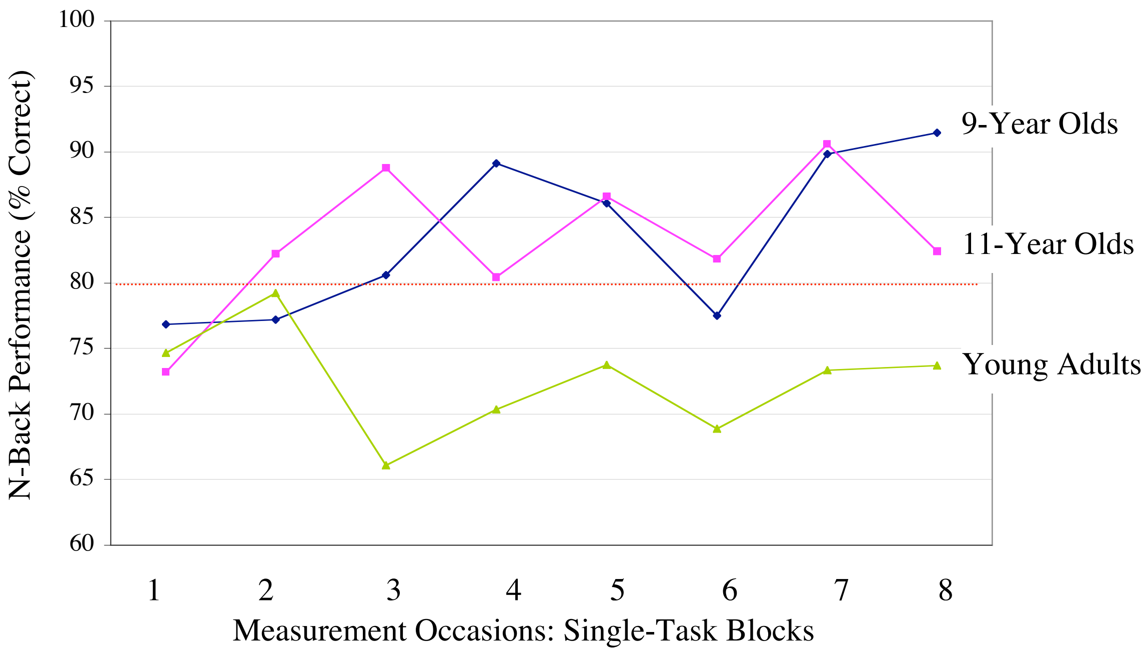

Figure 16. N-Back Performance Over the Course of Single-Task Assessment

Note. Performance is measured in percent correct. For each block, 2 single-task trials have been averaged, except for blocks 7 and 8, which represent the mean of 3 trials. Every dual-task session included 2 single-task blocks. The dotted red line represents the 80 % correct performance level.

Figure 16 depicts the average percent-correct performance on each of the single-task blocks of N-back. In a mixed-design ANOVA with single-task block (8) as within-subjects factor and age group (2) as between-subjects factor, neither the linear trend nor the quadratic trend reached significance in the analysis of the block effect, and these effects did not interact with the age contrasts. This indicates that the single-task baseline for N-back was stable over the course of the dual-task phase. However, on the level of overall performance, age contrasts revealed significant performance differences between children and young adults, $F(1,24) = 5.61$, $MSE = 1015.73$, $p < .05$, $\eta^2 = .190$, while the difference between 9- and 11-year-old children was not significant, $F(1,24) = .00$, $MSE = 1015.73$, $p = .953$, $\eta^2 = .000$. Figure 16 shows that from block 3 on, young adults perform consistently lower than children, which is another indicator that the N-back 4 version of the task was more difficult than the N-back 2 version.

5.3.3 Stability of Balance Baseline Performance

Does balance performance change systematically over the course of the dual-task phase, or does it remain stable? Tables I.1 and I.2 in Appendix I present the data for each trial separately, for each age group and the two different balance difficulty conditions. The data were analyzed statistically with four mixed-design ANOVAs with trial as the within-subjects factor and age group (2) as the between-subjects factor. Linear and quadratic trends were specified as contrasts for the within-subjects factor of trial. Analyses were conducted for each balance difficulty condition (stable and moving platform) and for each secondary task (reading numbers and detecting animal voices) separately. All four analyses showed that on the overall level, children had significantly larger COP areas than young adults, and 9-year olds had larger COP areas than 11-year olds.[15] The presentation of ANOVA results will therefore focus on the trial effect.

For the balance task on the stable platform while reading numbers, eight trials have been analyzed, because the single-task trials of the last session (differential-emphasis) were included in the analysis. The linear trend reached significance, $F(1,24) = 4.56$, $MSE = 11.44$, $p < .05$, $\eta^2 = .160$, and it did not interact with the age contrasts. The quadratic trend did not reach significance, and there was no interaction of this trend with the age contrasts. This indicates that balance performance in that condition improved in the last trials of the study. The analysis for the stable platform with detecting animal voices as the secondary task did not reveal a linear or a quadratic trend in the data, but the linear trend interacted with the age contrast comparing 9- to 11-year olds, $F(1,24) = 5.08$, $MSE = 21.21$, $p < .05$, $\eta^2 = .175$. A similar pattern emerged for the analysis of the moving platform while reading numbers: neither the linear nor the quadratic trend reached significance, but the linear trend interacted with the age contrast comparing 9- to 11-year olds, $F(1,24) = 4.49$, $MSE = 25.67$, $p < .05$, $\eta^2 = .158$. For the analysis of balance performance on the moving platform while detecting animal voices, neither the linear or the quadratic trend or any interaction of the age contrasts with theses trends reached significance. Taken together, no systematic change in balance performance over the course of the dual-task phase could be detected.

Gender differences in balance performance. The balance data were also influenced by gender, with girls and women showing consistently smaller COP areas than boys and men. For each of the 4 different single-task balance conditions, females swayed less than males, ($F(1,21) = 7.11$, $MSE = 5527.48$, $p < .05$ for balancing on the stable platform while reading

[15] There was only one exception in which the contrast comparing the balance performance of 9-year olds to 11-year olds did not reach significance, namely in analysis for the moving platform with detecting animal voices as the secondary task ($F(1,24) = 3.65$, $MSE = 164.04$, $p = .068$, $\eta^2 = .484$).

numbers; $F(1,21) = 10.13$, $MSE = 38924.91$, $p < .01$ for balancing on the stable platform while listening to animal voices; $F(1,21) = 10.71$, $MSE = 125726.57$, $p < .01$ for balancing on the moving platform while reading numbers, and $F(1,21) = 17.03$, $MSE = 274384.54$, $p < .001$ for balancing on the moving platform while listening to animal voices), and this effect did not interact with the effect of age group. The effect of gender did not disappear when height was controlled for in the analyses.

5.3.4 Summary of Results for the Training and Adaptive Phase of the Study and the Stability of Single-Task Baselines

Participants of all three age groups benefited equally from instruction in the MOL memory technique. Memory performance was further improved in the adaptive training phase, in which participants of each age group approached the 80 % correct performance level. The task difficulty levels which were finally chosen for the dual-task phase differed between age groups, with the young adults working on the longest lists, and the 9-year olds on the shortest lists. For the N-back task, a similar pattern emerged, with further improvements of task performance in the adaptive phase, and with an approach of the 80 % correct level in all three groups. The difficulty manipulation therefore successfully equated the performances of the three age groups in the two cognitive tasks.

In the balance task, age differences in balancing performance were apparent right from the start of the study, with young adults demonstrating smaller COP areas than children when working on comparable task conditions. Differences between 9-year olds and 11-year olds were not always significant. Task difficulty had a strong influence on balance performance: COP areas were larger on the moving platform as opposed to the stable platform in every age group. Furthermore, longer trials with detecting animal voices as secondary task resulted in larger COP areas than shorter trials with reading numbers as secondary task.

There were no systematic changes in single-task baseline performances of the three tasks across the dual-task phase of the study. The next section describes the performance in the dual-task phase of the study.

5.4 Dual-Task Performance

The following sections report the single- and dual-task performances for MOL, N-back and balance separately. Please note that the additional single-task trials that were assessed in the session with differential-emphasis instruction (Session 9) were also included in the single-task baselines for the N-back and balance task.

MOL. The MOL single-task baseline consisted of eight trials. For the dual-task performance, three trials for every balance difficulty condition were collected. Table J.1 in Appendix J presents the raw data for mean single- and dual-task performances for each age group, and Figure 17 depicts the MOL performance in percent correct under single-task conditions, when sitting in front of the computer, and under dual-task conditions, when standing on the stable platform or moving platform.

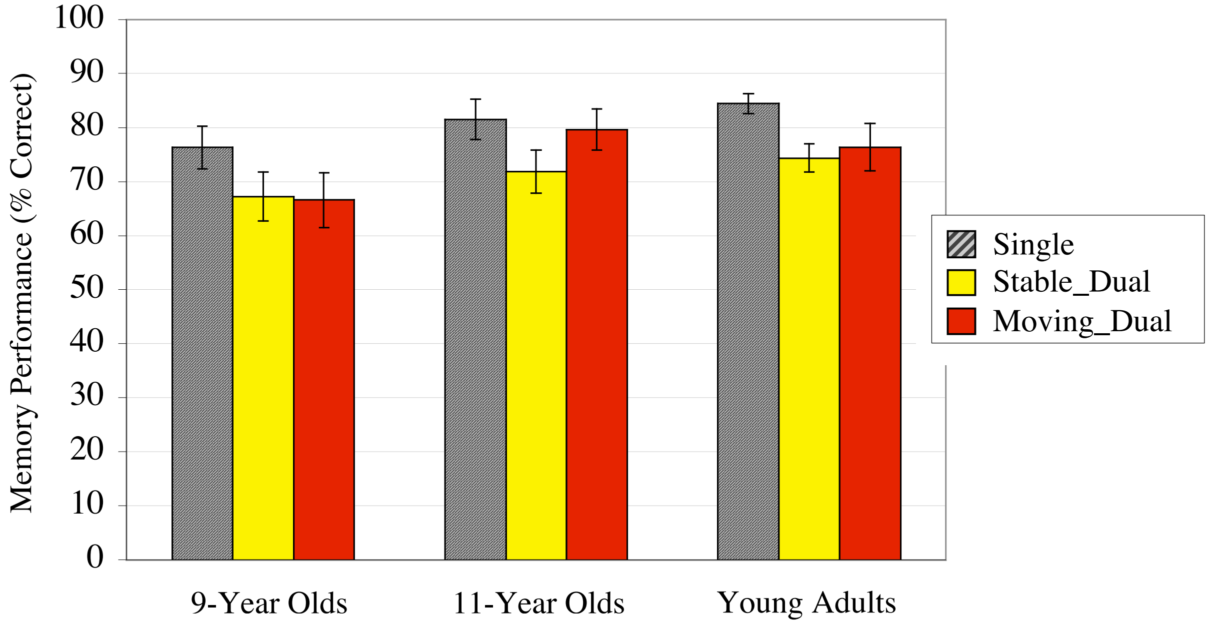

Figure 17. Method-of-Loci Performance Decreases Under Dual-Task Conditions

Note. Performance is measured in percent correct. The grey bars with the stripes (Single) depict performance under single-task conditions. The yellow bars (Stable_Dual) depict performance under dual-task conditions on the stable platform, and the red bars (Moving_Dual) on the moving platform. Error bars depict standard errors of the mean.

To investigate whether the performance reduction from single- to dual-task conditions was significant, a mixed-design ANOVA with task condition (3; single-task, dual-task on stable platform, and dual-task on moving platform) as within-subjects factor and with age group (2) as between-subjects factor on MOL percent correct scores as the dependent variable was conducted. For the effect of task condition, single-task performance was compared to the average of the dual-task performances, and the dual-task performance on the stable platform was compared to the dual-task performance on the moving platform. A significant difference was detected between performance in the single-task condition and performance in the average dual-task conditions, $F(1,24) = 25.81$, $MSE = 45.48$, $p < .01$, $\eta^2 = .518$, and that effect did not interact with the age contrasts, with $F(1,24) = .19$, $MSE = 45.48$, $p = .669$, $\eta^2 = .008$ for the contrast comparing children to adults, and $F(1,24) = .89$, $MSE = 45.48$, $p = .356$, $\eta^2 = .036$ for the contrast comparing 9- to 11-year olds. The comparison between the dual-task performance on the stable platform to the dual-task performance on the moving platform was not significant, $F(1,24) = 2.15$, $MSE = 58.49$, $p = .155$, $\eta^2 = .082$, and the interactions of the age contrasts with that effect were also not significant ($p = .740$ for children vs. adults and $p = .108$ for 9- vs. 11-year olds). In addition, the age contrast comparing children and young adults in their MOL performance did not reach significance in that analysis, $F(1,24) = 1.23$, $MSE = 302.05$, $p = .279$, $\eta^2 = .049$ and neither did the age contrast comparing 9- to 11-year old children, $F(1,24) = 2.59$, $MSE = 302.05$, $p = .120$, $\eta^2 = .097$. This indicates that the difficulty manipulation of the MOL task had equated the task performances of the three age groups. It can be concluded that MOL performance decreased under dual-task conditions, and this decrease was comparable for all three age groups. Furthermore, there was no difference between dual-task MOL performance on the stable as opposed to the moving platform.

N-back. For the N-back single-task baseline, the average percent correct performance consisted of 18 trials. The dual-task performance was assessed with 4 trials in each balance difficulty condition, and it required performing the N-back task while concurrently balancing on the stable or on the moving platform. Table J.2 in Appendix J presents the means and standard deviations of the N-back raw scores, and Figure 18 depicts the percent correct performance in N-back under single-task conditions (sitting in front of the computer) and under dual-task conditions on the stable or on the moving platform.

A mixed-design ANOVA on the N-back percent correct scores with task condition (3; single-task, dual-task on stable platform, and dual-task on moving platform) as within-subjects factor and with age group (2) as between-subjects factor was conducted. As in the analysis of MOL performance, the contrast for the task condition effect compared the single-task performance to the average of dual-task performances, and the dual-task performance on the stable platform to the dual-task performance on the moving platform. The difference between the single-task performance to the average dual-task performances reached significance, $F(1,24) = 7.66$, $MSE = 107.06$, $p < .05$. $\eta^2 = .242$. There was no interaction of this contrast with the age contrasts, $F(1,24) = .35$, $MSE = 107.06$, $p =.558$, $\eta^2 = .015$ for children versus adults, and $F(1,24) = 1.48$, $MSE = 107.06$, $p = .236$, $\eta^2 = .058$, for 9- versus 11-year olds. That indicates that performance in all three age groups decreased similarly under dual-task conditions. Participants tended to name fewer correct items when they concurrently had to balance on the ankle-disc board. In addition, there was a significant difference between dual-task trials on the stable as opposed to the moving platform, $F(1,24) = 7.23$, $MSE = 216.24$, $p < .05$, $\eta^2 = .231$, and that effect did not interact with the age contrasts, with $F(1,24) = 1.26$, $MSE = 216.24$, $p = .272$, $\eta^2 = .050$ for children versus adults, and $F(1,24) = 1.41$, $MSE = 216.24$, $p = .247$, $\eta^2 = .055$ for 9- versus 11-year olds. In all three age groups, N-back performance decreased further in the dual-task situation when the platform was moving instead of stable. Concerning the age contrast on an overall performance level, differences between the children and young adults were not significant, $F(1,24) = 3.83$, $MSE = 408.25$, $p = .062$, $\eta^2 = .138$, and neither were differences between 9- and 11-year olds, $F(1,24) = 1.03$, $MSE = 408.25$, $p = .320$, $\eta^2 = .041$. This shows that the difficulty manipulation led to comparable performance levels in the three age groups.

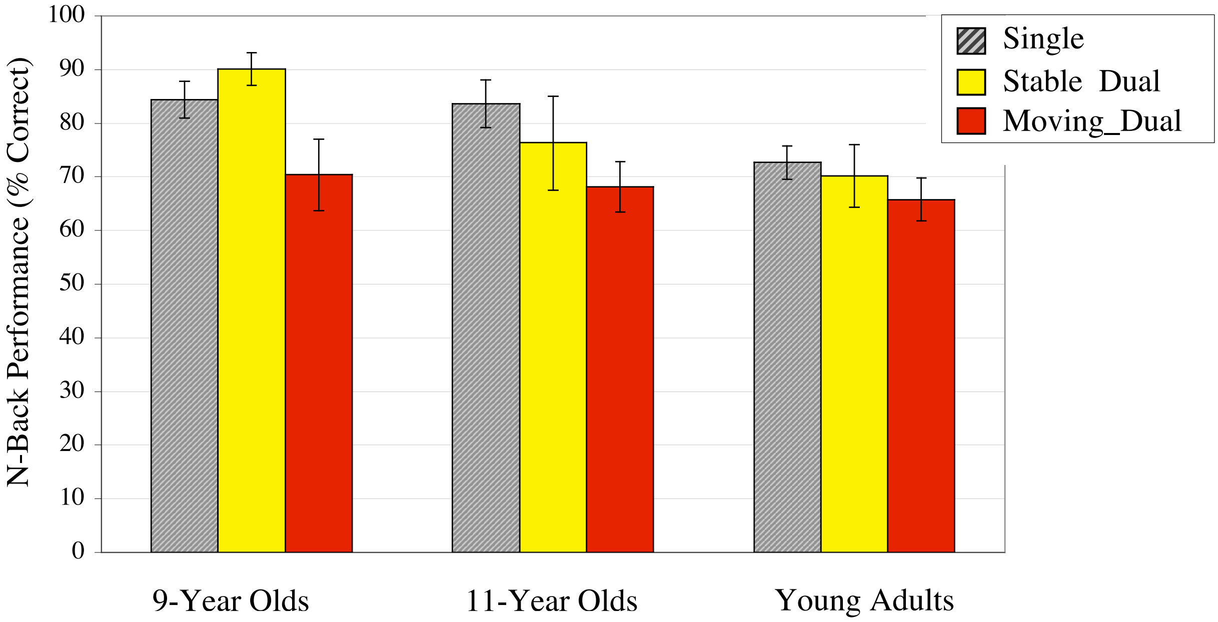

Figure 18. N-Back Performance Decreases Under Dual-Task Conditions, and the Decrease Is Influenced by Balance Difficulty

Note. Performance is measured in percent correct. The grey bars with the stripes (Single) depict performance under single-task conditions. The yellow bars (Stable_Dual) depict performance under dual-task conditions on the stable platform, and the red bars (Moving_Dual) on the moving platform. Error bars depict standard errors of the mean.

Balance. For the balance task, different “single-task” baseline conditions were used, with different secondary tasks for N-back (reading numbers) and MOL (detecting animal voices). Additionally, there were different balance difficulty conditions, namely the stable platform and the moving platform. Table J.3 in Appendix J presents the means and standard deviations of the COP areas under single- and dual-task conditions, with all possible combinations of secondary tasks (detecting animal voices and MOL, reading numbers and N-back). Results are displayed graphically in Figure 19 for balance with MOL, and in Figure 20 for balance with N-back. The most striking result was that performing a rather difficult cognitive task (N-back or MOL) concurrently while balancing did not lead to larger COP areas in children. Instead, the children's groups tended to show smaller COP areas than under single-task conditions.

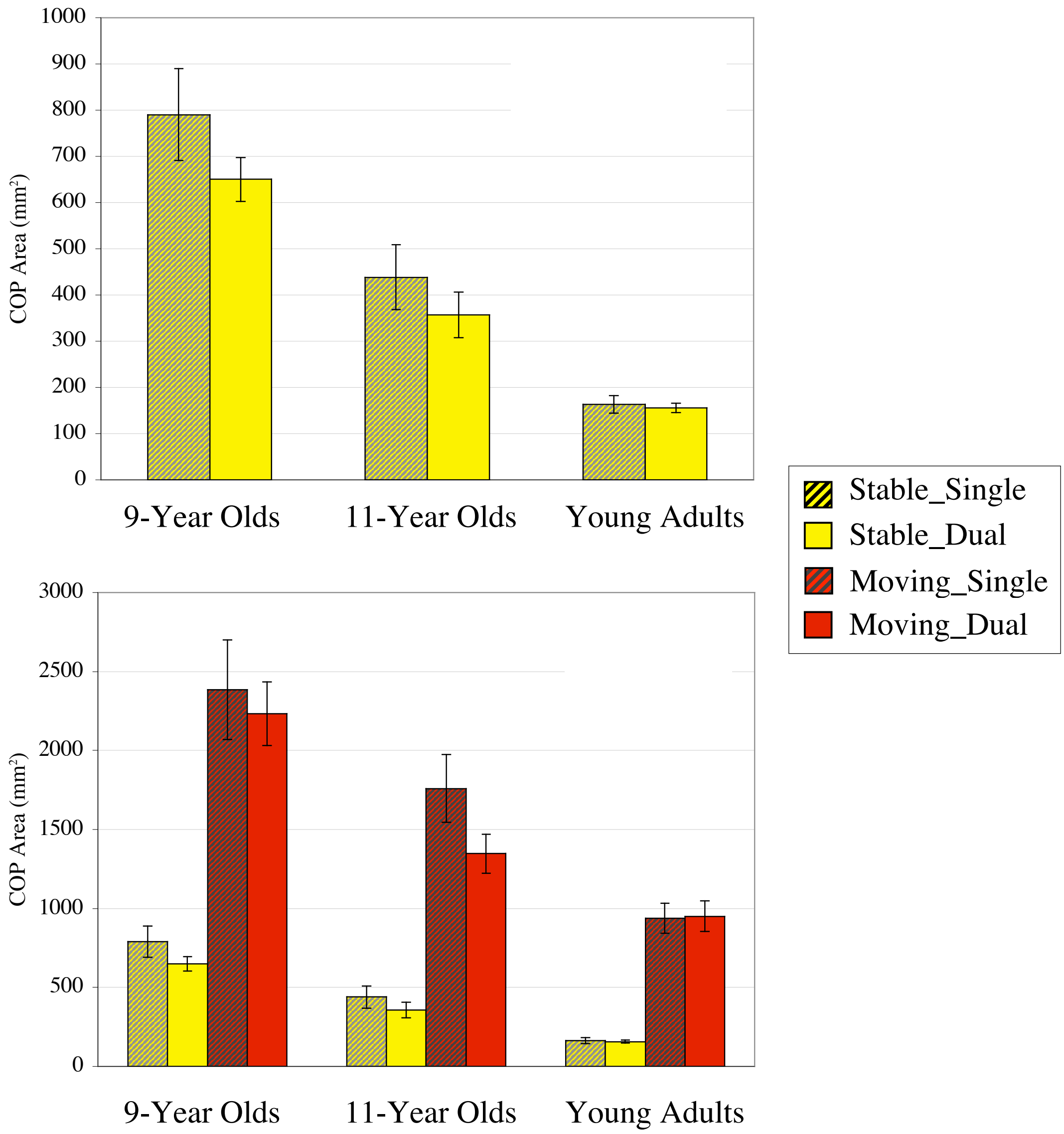

Figure 19. Single- and Dual-Task Balance Performance with Method-of-Loci as Cognitive Task

Note. The upper figure depicts performance on the stable platform (yellow bars), and the lower figure on the stable and moving platform (red bars). Striped bars represent balance performance under single-task conditions (when concurrently listening to animal voices), and solid bars under dual-task conditions (when concurrently performing the MOL task). Error bars depict standard errors of the mean.

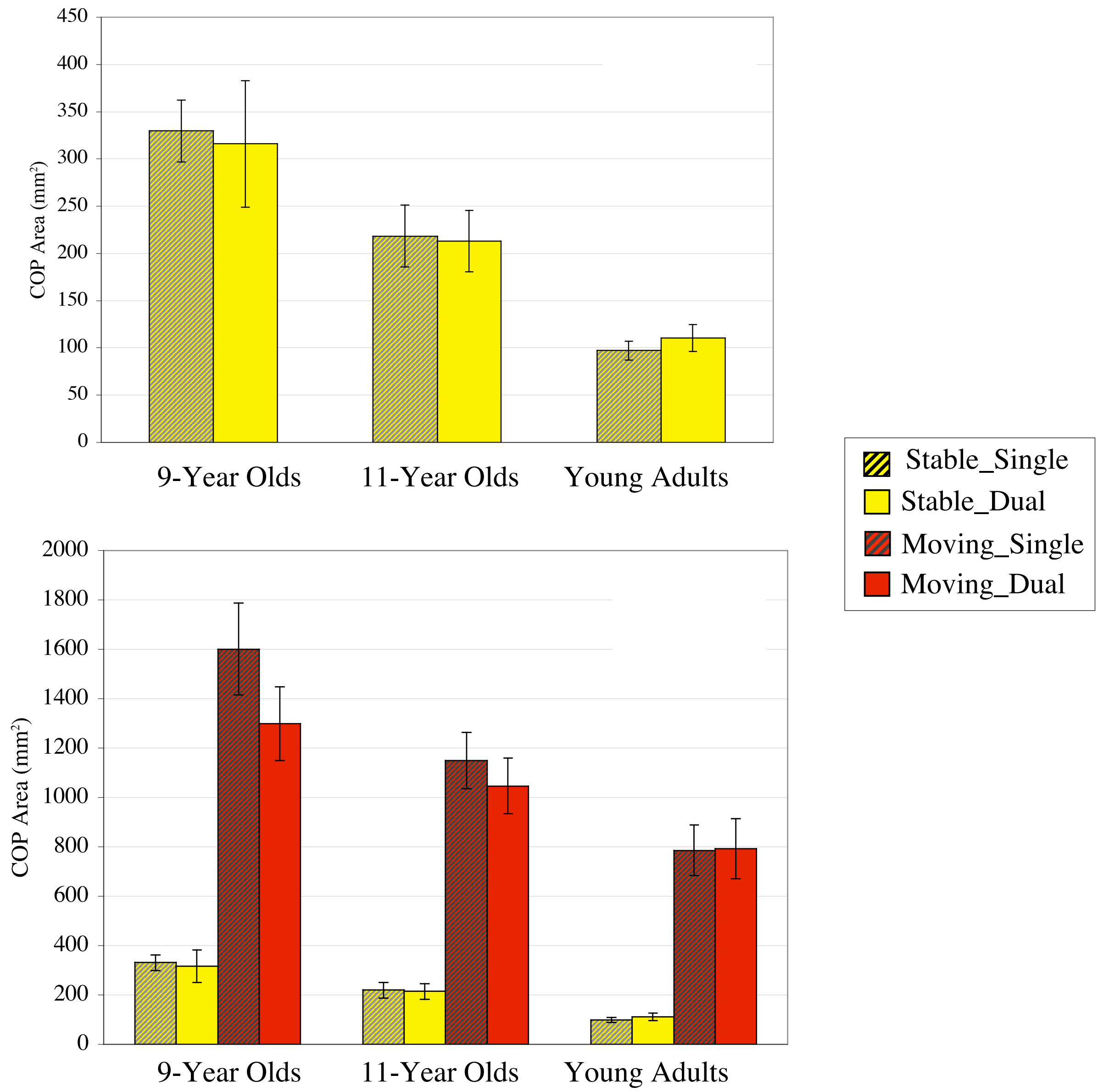

Figure 20. Single- and Dual-Task Balance Performance with N-Back as Cognitive Task

Note. The upper figure depicts performance on the stable platform (yellow bars), and the lower figure on the stable and moving platform (red bars). Striped bars represent balance performance under single-task conditions (when concurrently reading numbers), and solid bars under dual-task conditions (when concurrently performing the N-back task). Error bars depict standard errors of the mean.

Two mixed-design ANOVAs with single- versus dual-task (2) as within-subjects factor and age group (2) as between-subjects factor were conducted, for each balance difficulty (stable and moving platform) and each cognitive task (MOL and N-back) separately.[16] The results of the analysis for balance with MOL will be presented first.

For detecting animal voices versus performing the MOL task on the stable platform, the analysis revealed a significant difference between single- and dual-task performance, $F(1,24) = 4.74$, $MSE = 20391.47$, $p < .05$, $\eta^2 = .165$, and there was no interaction of this effect with the age contrasts, with $F(1,24) = 1.97$, $MSE = 20391.47$, $p = .174$, $\eta^2 = .076$ for children versus adults, and $F(1,24) = .13$, $MSE = 20391.47$, $p = .717$, $\eta^2 = .006$ for 9- versus 11-year olds. Figure 19 shows that COP areas were actually smaller under the dual-task as opposed to the single-task condition. Although the figure might indicate that this difference was more pronounced in the children than in the adults, this interaction did not reach significance in the analysis. Not surprisingly, young adults differed significantly from the children in their overall COP areas, $F(1,24) = 40.49$, $MSE = 48668.35$, $p < .001$, $\eta^2 = .628$, just as 9-year olds differed significantly from 11-year olds, $F(1,24) = 17.80$, $MSE = 48668.35$, $p < .001$, $\eta^2 = .426$. The same analysis was conducted for balance performance with MOL while balancing on the moving platform. Under this condition, the within-subjects effect of single- versus dual-task did not reach significance, $F(1,24) = 4.03$, $MSE = 113485.82$, $p = .056$, $\eta^2 = .144$, and it did not interact with the age contrasts. Age groups differed in their COP areas, with adults showing smaller COP areas than children, $F(1,24) = 21.24$, $MSE = 548340.97$, $p < .01$, $\eta^2 = .469$, and 11-year olds showing smaller COP areas than 9-year olds, $F(1,24) = 9.37$, $MSE = 548340.97$, $p < .01$. $\eta^2 = .281$.

The following section describes the results for the analysis of balance performance in combination with N-back as the cognitive task, again for each balance difficulty condition separately. For the stable platform, the mixed-design ANOVA revealed no significant differences between balancing while reading numbers versus working on the N-back task, $F(1,24) = .01$, $MSE = 6938.7$, $p = .920$, $\eta^2 = .000$, and there were no interactions of this effect with the age contrasts. As in the previous analysis, the age contrast comparing children to young adults was significant, $F\ (1,24) = 19.03$, $MSE = 17881.18$, $p < .01$, $\eta^2 = .442$, and 9-year olds differed significantly from 11-year olds, $F(1,24) = 6.45$, $MSE = 17881.18$, $p < .05$, $\eta^2 = .212$. On the moving platform, a significant within-subjects effect of single- versus dual-task was detected, $F(1,24) = 7.63$, $MSE = 31014.34$, $p < .05$, $\eta^2 = .241$. As can be seen in Figure 19, there was a reduction of COP areas from single- to dual-task conditions in the

[16] For the data in Table J.3 and for statistical analyses, outliers (see Table 8) have been adjusted by replacing their value with the second largest value of the respective distribution. However, running the analyses without adjusting for outliers did not change the pattern of results.

children, but not in the young adults, and this interaction almost reached significance, $F(1, 24) = 4.22$, $MSE = 31014.34$, $p = .051$, $\eta^2 = .150$, whereas the interaction of the single- versus dual-task effect with the age contrast comparing 9- to 11-year old children was not significant, $F(1,24) = 2.91$, $MSE = 31014.34$, $p = .101$, $\eta^2 = .108$. Overall, children again showed significantly larger COP areas than young adults, $F(1,24) = 9.53$, $MSE = 294628.02$, $p < .01$, $\eta^2 = .284$, but the difference between 9- and 11-year old children did not reach significance in this case, $F(1,24) = 3.79$, $MSE = 294628.02$, $p = .063$, $\eta^2 = .136$.

To summarize, it can be stated that the difficulty manipulation of the cognitive tasks led to comparable performance levels across age groups. In the MOL task, performance decreased under dual-task task conditions, but there was no influence of balance difficulty on the performance decrement. In the N-back task, performance also decreased under dual-task conditions, and it decreased more strongly on the moving as opposed to the stable platform. In both cognitive tasks, age groups did not differ in their performance reduction under dual-task conditions. For the balance task, the dual-task situation actually led to performance improvements in various conditions, and these improvements were more pronounced in children than in adults.

Appendix K presents correlations of cognitive and sensorimotor measures of the present study with measures assessed in the screening session. In addition, Appendix L presents age differences in balance performance when each single-task balance score is expressed in percent of individual FSB area used.

5.4.1 Calculation of Dual-Task Costs

Performance reduction from single-to dual-task situations can be expressed as *dual-task costs* (DTCs). For *absolute DTCs*, performance reductions are expressed in the same metric as the task performance. For example, if a participant remembers 15 MOL items under single-task conditions and only 12 items under dual-task condition, the absolute dual-task costs would be 3 items. However, another participant who only remembers 7 items under single-task conditions and 4 items under dual-task conditions has the same value for absolute dual-task costs, namely 3 items, although the overall performance level differs considerably between the two participants. Absolute DTCs can therefore be misleading when interpreting performance patterns of dual-task studies.

To compare the performance changes from single- to dual-task performance across different domains (e.g., cognition and balance) and cognitive tasks (e.g., MOL and N-back) in the present study, *proportional DTCs* were calculated as suggested by Somberg and Salthouse

(1982). Proportional DTCs express the performance reduction in percent of single-task performance, by taking each individual's baseline performance level into account. Proportional DTCs for the cognitive tasks (N-back and MOL), in which better performance is represented by higher scores, were calculated according to:

[(dual-task performance – single-task performance) / single-task performance] * -100 %

For the above MOL example in which a participant remembers 15 items under single-task conditions and 12 under dual-task conditions, the calculation would be the following:
[(12–15) / 15] *-100 % = 20 %.
That result would be interpreted as 20 % performance reduction in MOL caused by the simultaneous performance of the balance task. For balance, in which better performance is represented by smaller COP areas (and therefore in lower scores), DTCs were calculated according to:

[(dual-task performance – single-task performance) / single-task performance] * 100 %

Expressed as proportional DTCs, performance improvements under dual-task conditions lead to *negative* dual-task costs.

Figure 21 gives an overview of the proportional DTCs for each task and balance difficulty condition separately, with cognitive DTCs in the upper half and balance DTCs in the lower half of the figure. Costs in cognition were positive and reliably different from zero in 8 out of 12 instances (age group x conditions), indicating that the cognitive performance was reduced when the balance task had to be performed concurrently. In contrast, balance dual-task costs were negative in the children in 6 out of 8 conditions, indicating that children were able to improve their balance performance under dual-task conditions. Young adults, on the other hand, generally showed positive costs (performance decrements) in the balance domain, which were significantly different from zero only in one condition.

Table J.4 in Appendix J presents the DTCs for MOL, N-back and balance, for the balance difficulty conditions of stable and moving platform. Outliers were again replaced by the second largest value in the respective distribution.

The following sections report the data patterns concerning the hypotheses of the study.

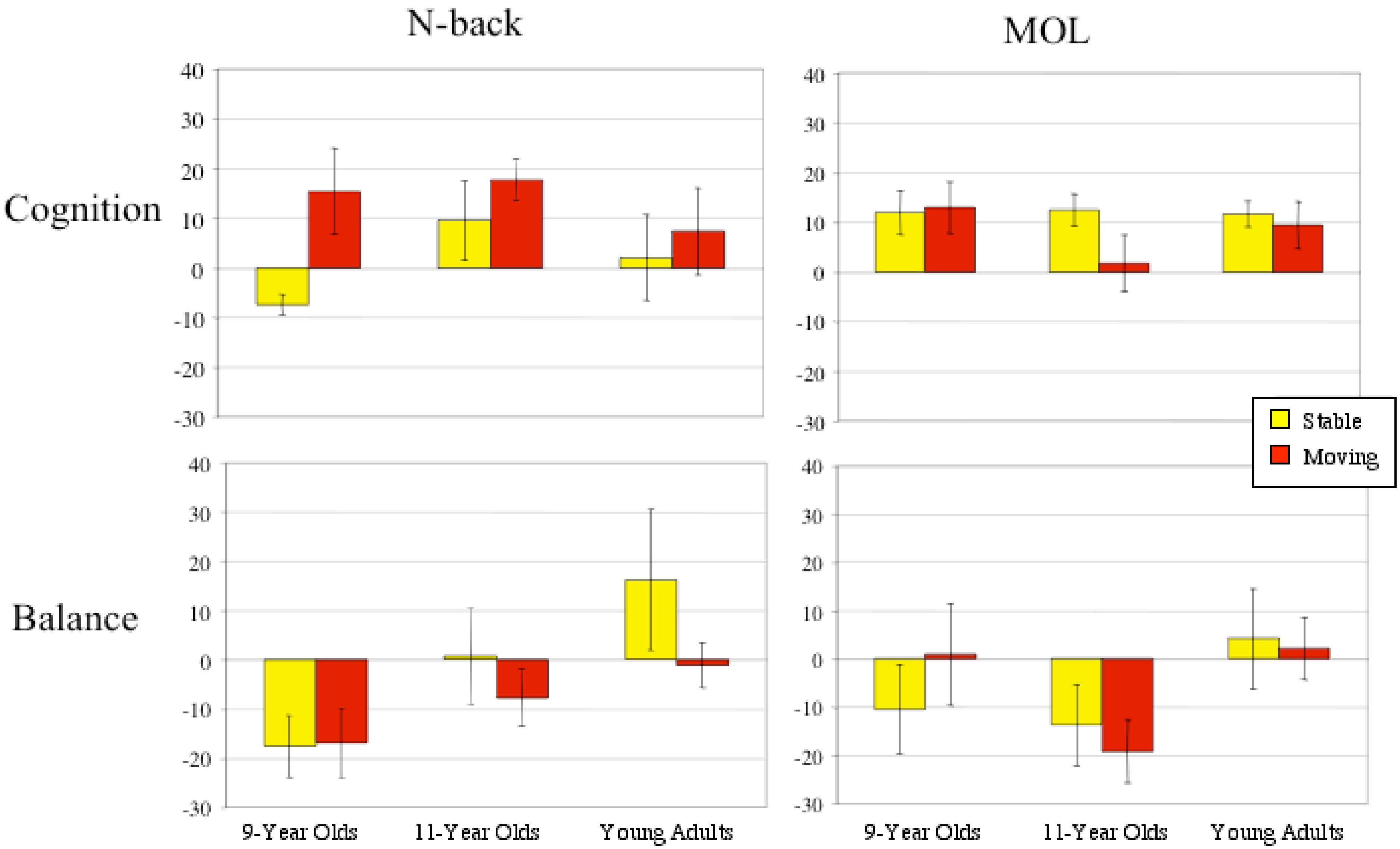

Figure 21. Overview of Dual-Task Costs by Task Modality, Cognitive Task, and Balance Difficulty

Note. The columns stand for dual-task costs (DTCs) of the two different cognitive tasks: the left column for the N-back task, and the right column for the MOL task. The different rows represent the task domains: the cognitive DTCs are depicted in the upper half, and the balance DTCs in the lower half of the figure. Within each figure, DTCs on the stable platform are represented by the yellow bars and DTCs on the moving platform by the red bars. Error bars depict standard errors of the mean.

5.4.2 Age Differences in the Trade-Off Pattern Between Cognitive and Balance DTCs

Since proportional DTCs use the same metric across tasks, it is possible to directly compare the dual-task costs of different task domains assessed under different difficulty conditions. Figure 22 presents the cognitive and balance DTCs of the present study, which were aggregated over balance task difficulty (stable and moving platform) and cognitive task (MOL and N-back). Additionally, the DTCs of the two children's groups were aggregated.

To investigate whether there are differences in the DTCs between the two task modalities of cognition and balance, between the two cognitive tasks (N-back and MOL), or between the two balance difficulty conditions (stable and moving platform), a mixed-design ANOVA has been conducted. The following DTCs were analyzed as dependent variables in the ANOVA: balance DTCs with N-back and MOL on the stable and the moving platform, and cognitive DTCs for N-back and MOL on the stable and the moving platform, resulting in 8 DTC-measures. For these dependent variables, the within-subjects effects of task modality (2; cognition vs. balance), cognitive task (2; N-back vs. MOL), and balance difficulty (2; stable vs. moving) were investigated. Age group (2) was the between-subjects factor, by comparing 9-year olds to 11-year olds, and the two children's groups to the young adults.

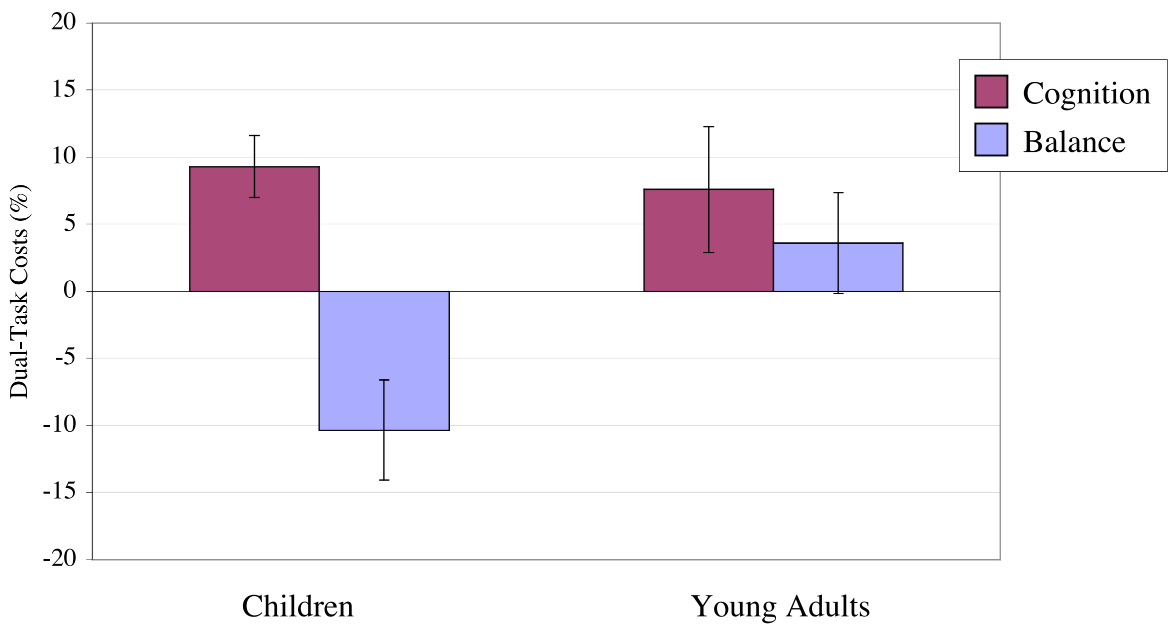

Figure 22. Children Show a Trade-Off Pattern Between Cognitive and Balance Dual-Task Costs, Young Adults Do Not

Note. Data for the two children's groups (N = 18) were averaged, and DTCs were averaged over balance difficulty and cognitive task. Error bars depict standard errors of the mean.

The analysis of the between-subjects effect of age group revealed no significant differences between children and young adults, $F(1,24) = 2.42$, $MSE = 742.61$, $p = .133$, $\eta^2 = .092$, or between 9- and 11-year olds, $F(1,24) = .14$, $MSE = 742.61$, $p = .715$, $\eta^2 = .006$. This means that the overall DTCs (averaged over task modality, cognitive task, and balance difficulty) did not differ between the age groups.

For the within-subjects effect of *task modality*, a significant main effect was detected, $F(1,24) = 16.38$, $MSE = 685.46$, $p < .001$, $\eta^2 = .406$, and there was no significant interaction of the modality effect and the age contrast comparing 9- to 11-year olds, $F(1,24) = .01$, $MSE = 685.42$, $p = .917$, $\eta^2 = .000$, whereas there was a marginally significant interaction of the modality effect and the age contrast comparing children to young adults, $F(1,24) = 4.27$, $MSE = 685.42$, $p = .050$, $\eta^2 = .151$. This indicates that differences between cognitive and balance dual-task costs existed in the children. The data pattern therefore supported the first hypothesis of the study, namely that children invest more of their cognitive resources into the sensorimotor domain and prioritize balance performance over cognitive performance when the body's equilibrium is challenged.

The within-subjects main effect of *cognitive task* was not significant, $F(1,24) = .15$, $MSE = 439.99$, $p = .703$, $\eta^2 = .006$, but there was a significant interaction of that effect with the age contrast comparing 9- to 11-year olds, $F(1,24) = 8.11$, $MSE = 439.99$, $p < .05$, $\eta^2 = .253$, and no interaction with the contrast comparing children to young adults, $F(1,24) = .11$, $MSE = 439.99$, $p = .744$, $\eta^2 = .005$. This means that there were little differences between all DTCs with the N-back task and all DTCs with the MOL task (averaged over task modality and balance difficulty), but that 9-year olds differed considerably from 11-year olds in these differences.

Furthermore, the within-subjects effect of *balance difficulty* was not significant, $F(1,24) = .05$, $MSE = 578.94$, $p = .818$, $\eta^2 = .002$, and it did not interact with any of the age contrasts (9-year olds vs. 11-year olds with $p = .110$, and children vs. adults with $p = .516$). None of the possible two-way interactions between within-subjects factors reached significance (task modality * cognitive task: $p = .574$, task modality * balance difficulty: $p = .226$, balance difficulty * cognitive task: $p = .351$), and neither did the respective age contrasts of each two-way interaction. However, the three-way interaction between task modality, cognitive task and balance difficulty was significant, $F(1,24) = 18.39$, $MSE = 98.50$, $p < .001$, $\eta^2 = .434$, but this effect did not interact with the age contrasts (9-year olds vs. 11-year olds with *p* = .389, and children vs. adults with *p* = .351). This was caused by the influence of balance difficulty on cognitive DTCs in the N-back task, but not in the MOL task.[17]

To follow up the ANOVA results, several t-tests were conducted. For a closer investigation of the significant main effect of task modality (cognition versus balance), children's DTCs in cognition were compared to young adults' DTCs in cognition, and there were no significant differences between these groups, $t(25) = .363$, $p = .720$. However, children's DTCs in balance differed significantly from young adults' DTCs in balance, $t(25) = -2.33$, $p < .05$. In addition, paired samples t-tests comparing cognitive DTCs to balance DTCs in all three age groups revealed significant differences between the two task modalities, $t(26) = 3.88$, $p < .05$. Analyzing this difference only for the two children's groups revealed significant differences between cognitive and balance DTCs, $t(17) = 4.31$, $p < .001$, but there were no significant differences between cognitive and balance DTCs in the young adults, $t(8) = .784$, $p = .456$. These results corroborated the finding that children showed a performance trade-off between cognition and balance, favoring the balance performance, whereas young adults did not show a trade-off pattern in their cognitive and balance dual-task costs.

The interaction of the effect of cognitive task (N-back vs. MOL) with the age contrast comparing 9- to 11-year olds was investigated in more detail by comparing the N-back DTCs of the 9-year olds to the N-back DTCs of the 11-year olds in an independent samples t-test.[18] The t-test did reveal significant differences between 9- and 11-year olds in N-back DTCs, $t(16) = -2.20$, $p < .05$, but there were no significant differences between the two children's groups in a t-test for MOL DTCs, $t(16) = 1.48$, $p = .159$. A paired samples t-test comparing N-back DTCs to the MOL DTCs for all three age groups did not detect significant differences between the two cognitive tasks, $t(26) = .347$, $p = .731$. However, when the same kind of test was conducted with only the 9-year olds' data, significant differences between N-back DTCs and MOL DTCs were found, $t(8) = 2.67$, $p < .05$, while the same analysis did not reveal significant differences in the 11-year olds, $t(8) = -1.87$, $p = .099$.

---

[17] Conducting the same analysis without adjusting the outliers yielded the same pattern of results, except that the interaction of the within-subjects effect of task modality did not interact with the age contrast comparing children and young adults, $F(1,24)= 3.09$, $p = .092$. The other exception was that the interaction of the within-subjects effect of cognitive task did not interact with the age contrast comparing 9- to 11-year olds, $F(1,24)= 3.90$, $p = .060$.

[18] Note that "N-back DTCs" in this context comprise all DTCs that were assessed in combination with the N-back task, averaging over balance difficulty and task modality.

Figure 22 and the results of the above analyses show that there was a pronounced difference between the DTCs of the two modalities in children, with higher DTCs in the cognitive domain than in the balance domain, whereas the difference was much smaller for young adults. Children's balance DTCs were negative, which indicates that their balance performance actually improved under dual-task conditions. This trade-off pattern in the children had been expected according to the first hypothesis of the study.

5.4.3 The Influence of Balance Difficulty on the Trade-Off Pattern

The second hypothesis proposed that the difference between cognitive and balance DTCs would be even more pronounced under the more difficult balance condition (when balancing on the moving platform instead of the stable platform) in the children, but not in the adults. Figure 23 depicts the cognitive and balance DTCs for children (averaged over 9- and 11-year olds) and young adults on the stable and on the moving platform.

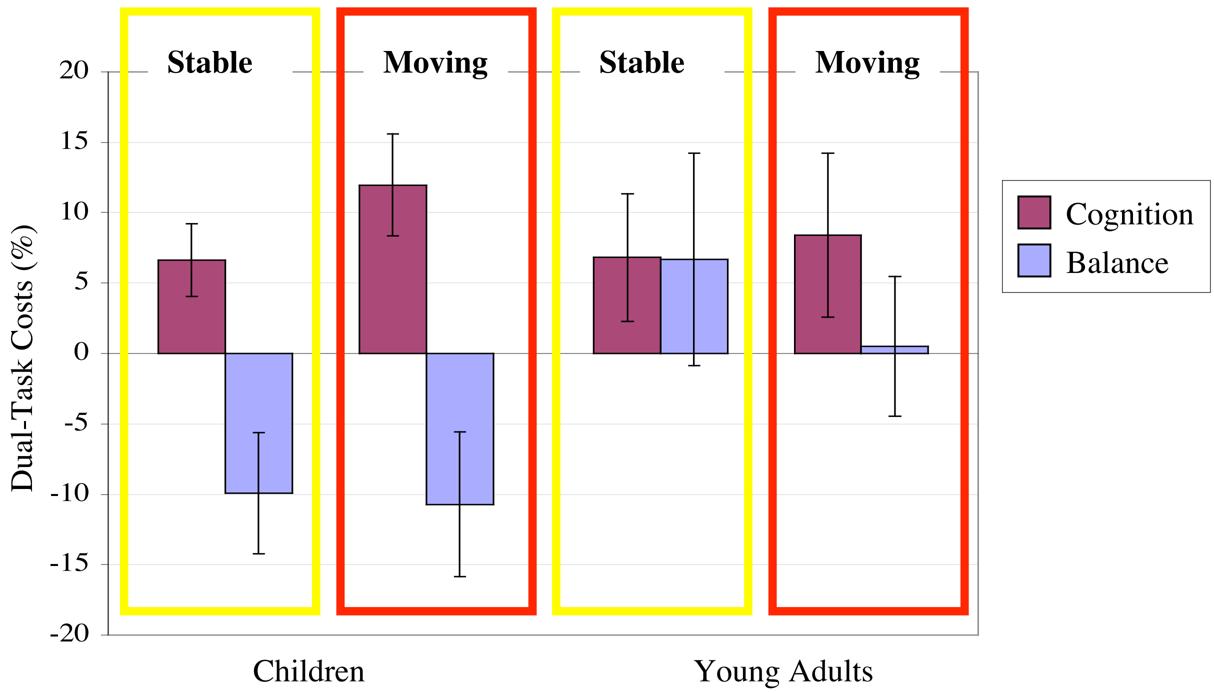

Figure 23. Children Do Not Show a More Pronounced Trade-Off Pattern When Balance Difficulty Is Increased

Note. DTCs were averaged over cognitive task (N-back and MOL), and data for the two children's groups (N = 18) were averaged. The yellow frame represents DTCs on the stable platform, and the red frame on the moving platform. Error bars depict standard errors of the mean.

It is obvious that the children show the trade-off between cognition and balance in both difficulty conditions. There is a tendency in the children to show more pronounced DTCs in the cognitive domain on the moving platform as compared to the stable platform. However, the mixed-design ANOVA presented in the previous section showed that there was no interaction of balance difficulty, task domain, and the age contrast comparing children to young adults ($F(1,24) = .02$, $MSE = 187.26$, $p = .887$, $\eta^2 = .001$).

To follow up the pattern of data, paired samples t-tests with the DTCs of each task domain (balance vs. cognition) on the stable and moving platform were conducted, for the children and the young adults separately. No significant differences between the cognitive DTCs on the stable versus the moving platform were found in the children, $t(17) = -1.26$, $p = .225$, or in the young adults, $t(8) = -.35$, $p = .735$. The balance DTCs also did not differ significantly on the stable versus the moving platform in the children, $t(17) = .14$, $p = .893$, or the young adults, $t(8) = .60$, $p = .565$.

It can be concluded that balance difficulty did not influence the dual-task costs in the current study, neither on an overall nor on a domain-specific level. Balance difficulty also did not increase the trade-off pattern between balance and cognition in the children, such that the second hypothesis was not supported by the data.

5.5 Differential-Emphasis Manipulation

Concerning the trade-off between the two task modalities, the question was whether this performance pattern was influenced by systematically attending more to one task than the other in a dual-task situation. The third hypothesis argued that children would not be able to shift their attention away from the balance domain, and would continue to show the trade-off pattern even when they were instructed to focus primarily on the cognitive task. In the last session of the study, a differential-emphasis instruction encouraged participants to focus more on one task than on the other. In that session, only N-back was used as the cognitive task, and all balance trials were performed on the stable platform. Three different emphasis instructions were used: “focus on the N-back task”, “focus on the balance task”, and “focus on both tasks equally well”. The following section does not report the findings for the condition to “focus on both tasks equally well”, because these data might blur the results of interest, namely the comparison of the conditions to focus on the N-back task versus to focus on the balance task.

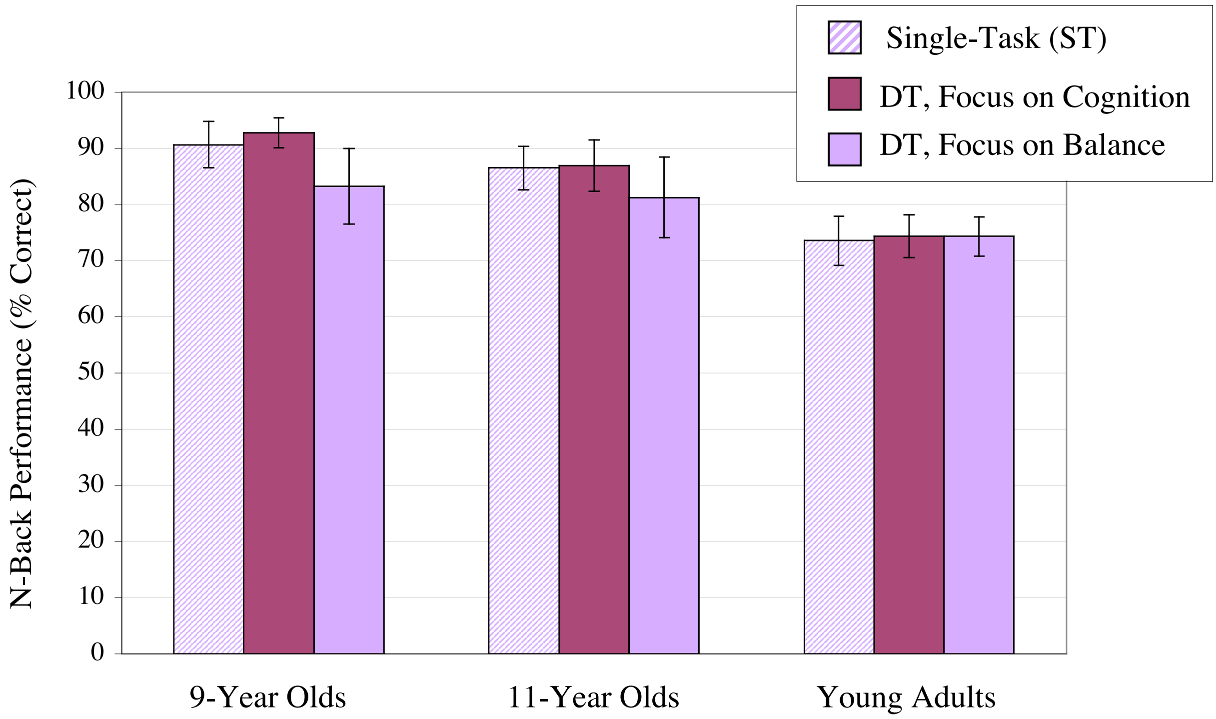

Figure 24. N-Back Performance Under Differential-Emphasis Instructions

Note. N-back performance is measured in percent correct. The striped bars depict single-task performance, the dark purple bars depict performance under the condition to "focus more on the N-back task", and the light purple bars under the condition to "focus more on the balance task". Error bars depict standard errors of the mean.

Raw Scores. Figure 24 presents the performance pattern for the N-back task in the differential-emphasis phase, and Figure 25 presents the performance pattern for the balance task. The single-task baselines were calculated by only using single-task trials of the differential-emphasis session.

A mixed-design ANOVA on the N-back percent-correct scores was conducted, with single- versus dual-task as a within-subjects factor (3; single-task, dual-task "focus on the N-back task", dual-task "focus on the balance task"). Age group (2) was used as the between-subjects factor. The contrast used for the single- versus dual-task factor compared each of the dual-task conditions ("focus on N-back" and "focus on balance") to the single-task condition. None of these contrasts reached significance, and there were no interactions of the effect with the age contrasts. In addition, children differed significantly from the young adults in their N-back performance, $F(1,24) = 7.56$, $MSE = 392.60$, $p < .05$, $\eta^2 = .240$, while 9-year olds did not differ from 11-year olds, $F(1,24) = .54$, $MSE = 392.60$, $p = .471$, $\eta^2 = .022$. This pattern indicates that the N-back percent-correct performance generally did not change from single-to dual-task, no matter what the differential-emphasis instruction indicated, and the differences between the single- and the dual-task condition did not differ by age group.

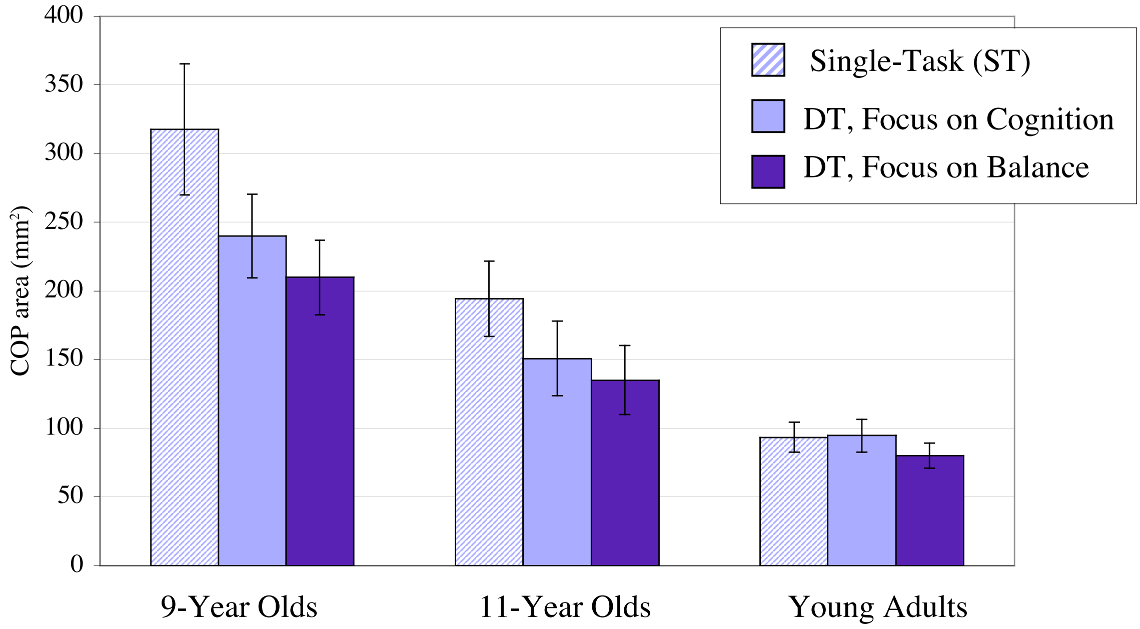

Figure 25. Balance Performance Under Differential-Emphasis Instructions

Note. The striped bars depict single-task performance, the light blue bars depict performance under the condition to "focus more on the N-back task", and the dark blue bars under the condition to "focus more on the balance task". Error bars depict standard errors of the mean.

Another mixed-design ANOVA was conducted for the COP areas of the balance task, with single- versus dual-task as within-subjects factor (3), and with age group (2) as between-subjects factor. For the within-subjects effect, a contrast comparing each of the two different emphasis-conditions to the single-task performance was used. The balance performance under the instruction to "focus on the N-back task" did not differ significantly from the single-task balance performance, $F(1,24) = .99$, $MSE = 1787.04$, $p = .330$, $\eta^2 = .040$, and this effect did not interact with the age contrasts. However, there was a significant difference between balance performance under the instruction to "focus on the balance task" and the single-task condition, $F(1,24) = 32.35$, $MSE = 1510.88$, $p < .01$, $\eta^2 = .574$, and this effect interacted with the age contrast comparing children to adults, $F(1,24) = 9.68$, $MSE = 1510.88$, $p < .01$, $\eta^2 = .287$, while the interaction with the age contrast comparing 9- to 11-year olds did not reach significance, $F(1,24) = 3.57$, $MSE = 1510.88$, $p = .071$, $\eta^2 = .129$. Figure 25 shows that the significant differences were caused by an improvement of balance performance in the dual-task as opposed to the single-task situation. Participants showed smaller COP areas under dual-task conditions when focusing on the balance task, and these differences were more pronounced in the children than in the young adults. On the overall performance level, just as in the other analyses with balance data, adults showed smaller COP areas than children, $F(1,24) = 15.89$, $MSE = 15907.64$, $p < .01$, $\eta^2 = .398$, and 11-year olds showed smaller COP areas than 9-year olds, $F(1,24) = 7.78$, $MSE = 15907.64$, $p < .05$, $\eta^2 = .245$. The next section investigates the performance pattern when the change in performance from single to dual-task is expressed in proportional dual-task costs.

Proportional dual-task costs. To investigate the differences between the two emphasis conditions in a proportional metric, DTCs for the differential-emphasis phase have been calculated, which express performance reductions (or improvements) in percent of single-task performance. Figure 26 depicts the resulting dual-task pattern for the children, comparing the two instruction conditions, and Figure 27 depicts the pattern for young adults. Figure 26 shows that the balance dual-task costs remain negative in the children, independent of task emphasis instruction.

A mixed-design ANOVA with emphasis-condition (2; "focus on N-back" vs. "focus on balance") and task domain (2, N-back vs. balance) as within-subjects factors and age group (2) as between-subjects factor was conducted. The main effect of emphasis-condition was not significant, $F(1,24) = 1.12$, $MSE = 189.91$, $p = .301$, $\eta^2 = .045$, and it did not interact with the age contrasts. This means that the overall DTCs (averaged across task modality) did not differ by emphasis-condition, and age groups did not differ in this pattern. The main effect of modality was significant, $F(1,24) = 14.69$, $MSE = 630.18$, $p < .01$, $\eta^2 = .380$, and there was an interaction of the modality effect and the age contrast comparing children to adults, $F(2,24) = 8.28$, $MSE = 630.18$, $p < .01$, $\eta^2 = .256$, but not with the age contrast comparing 9- to 11-year olds, $F(1,24) = .10$, $MSE = 630.18$, $p = .754$, $\eta^2 = .004$. This indicates that DTCs in the balance domain were different from DTCs in the cognitive domain, and children showed more pronounced differences between the two domains than adults, with smaller DTCs in the balance domain than in the cognitive domain. In addition, there was a significant interaction between task modality and emphasis condition, $F(1,24) = 10.24$, $MSE = 222.29$, $p < .01$, $\eta^2 = .299$, and this interaction did not interact with the age contrasts, showing that the difference between balance DTCs and cognitive DTCs also depended on emphasis condition. There were larger differences between balance and cognitive DTCs under the instruction to "focus on the balance task", but that effect was comparable across age groups.

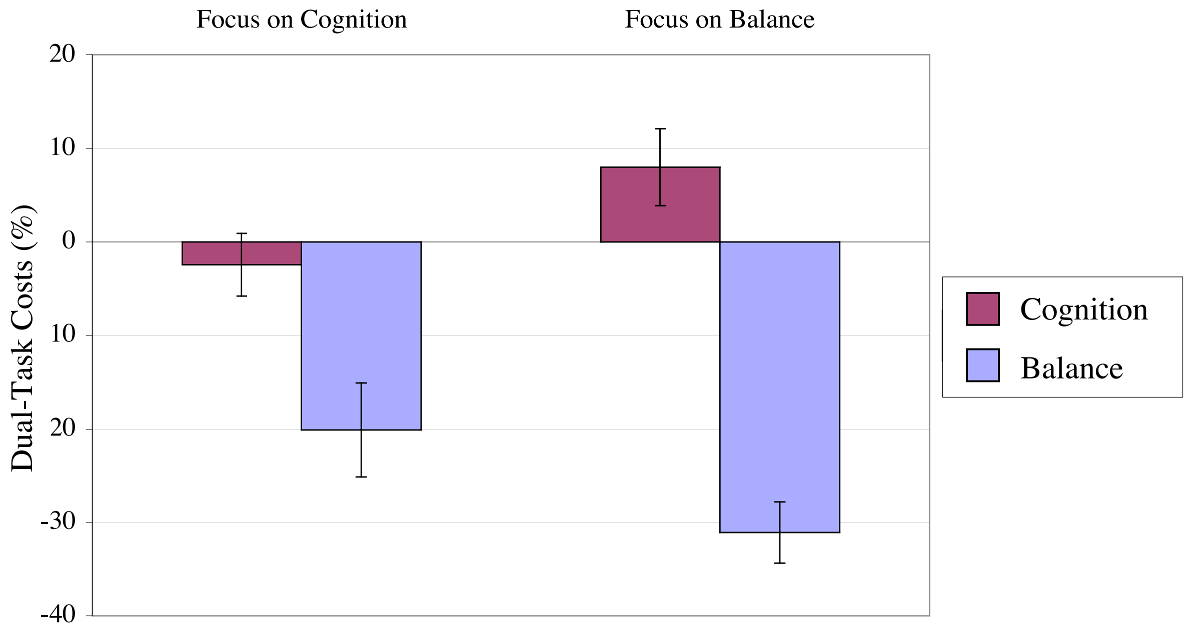

Figure 26. Children Continue to Show a Trade-Off Pattern in Their Dual-Task Costs Under Differential-Emphasis Instructions

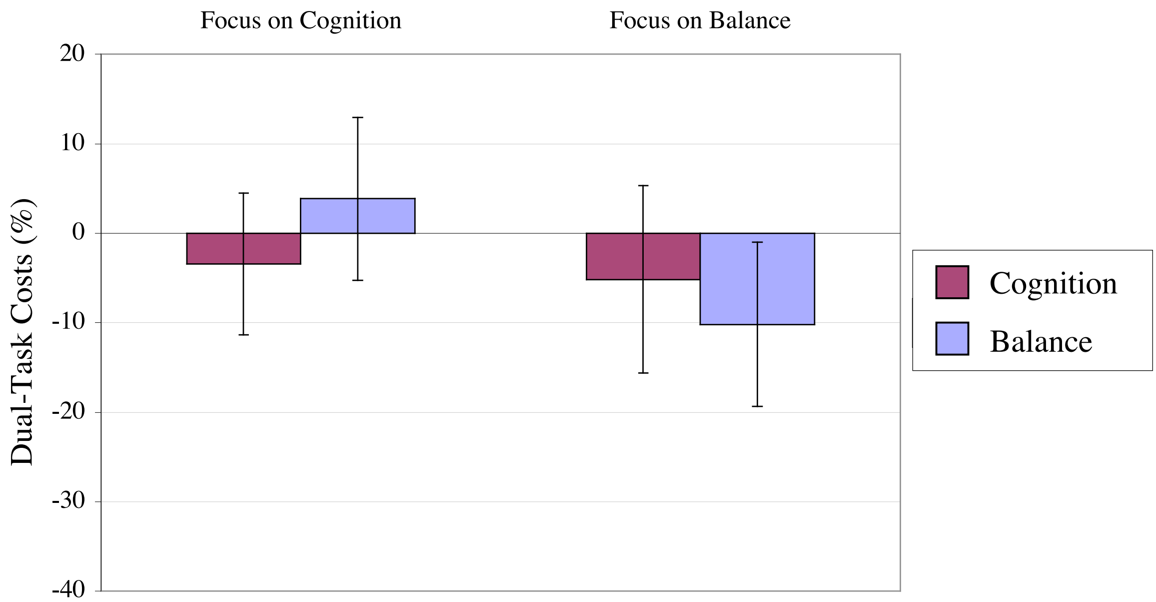

Figure 27. Young Adults' Dual-Task Costs Do Not Differ By Instruction in the Differential-Emphasis Phase

Note. Data of the two children's groups in Figure 26 were averaged. Purple bars present cognitive DTCs, and blue bars present balance DTCs. Error bars depict standard errors of the mean.

For the overall DTCs, children did not differ significantly from young adults, $F(1,24) = 1.79$, $MSE = 788.16$, $p = .193$, $\eta^2 = .069$, and 9-year olds did not differ from 11-year olds, $F(1,24) = .03$, $MSE = 788.16$, $p = .866$, $\eta^2 = .001$. The overall DTCs were generally very low for that session (9-year olds: $M = -10.85$, $SD = 8.69$; 11-year olds: $M = -11.98$, $SD = 9.71$; young adults: $M = -3.74$, $SD = 20.53$). This indicates that the reinforcement procedure motivated participants in all three age groups to perform at a high level.

The third hypothesis claimed that children tend to not shift their attention away from balance, even when instructed to do so, while young adults can allocate their resources according to instruction. To investigate the data concerning that hypothesis in further detail, paired-samples t-tests were calculated. These tests compared DTCs in the balance domain to DTCs in the cognitive domain, for each emphasis condition separately. Data of the two children's groups have been aggregated, since their performance levels and DTCs did not differ significantly, and there were no interactions of the within-subjects effects with the age contrast comparing 9- to 11-year olds in the analyses presented above.

For the condition "focus on N-back", significant differences between the DTCs of the two domains were detected in the children, $t(17) = -2.77$, $p < .05$, but not in the young adults, $t(8) = .651$, $p = .533$. A similar pattern emerged for the condition "focus on balance", again with significant differences between cognitive and balance DTCs in the children, $t(17) = -6.87$, $p < .001$, but not in the young adults, $t(8) = -.428$, $p = .680$. In all theses cases, the significant results were due to the fact that children showed consistently lower DTCs in the balance task as opposed to the N-back task. Therefore, the hypothesis that children have a harder time than young adults to shift their attention away from balance was supported by the data. Young adults, however, did not show differences in the DTCs of the two domains under any emphasis condition.

Reinforcement points. An alternative way to investigate the processes of resource allocation in the differential-emphasis phase is to analyze the distribution of reinforcement points that participants gained when reaching the pre-specified performance criterion of a certain task domain. It should be noted that the calculation of performance criteria for the differential-emphasis session took each individual's performance from the previous dual-task session with the same task parameters into account. This insured that people who performed at very high levels already in previous dual-task sessions had to improve their performance even further to collect points. Figure 28 depicts the distribution of reinforcement points by age group, emphasis condition, and task domain as the proportion of potentially achievable points that actually were reached in a specific condition.

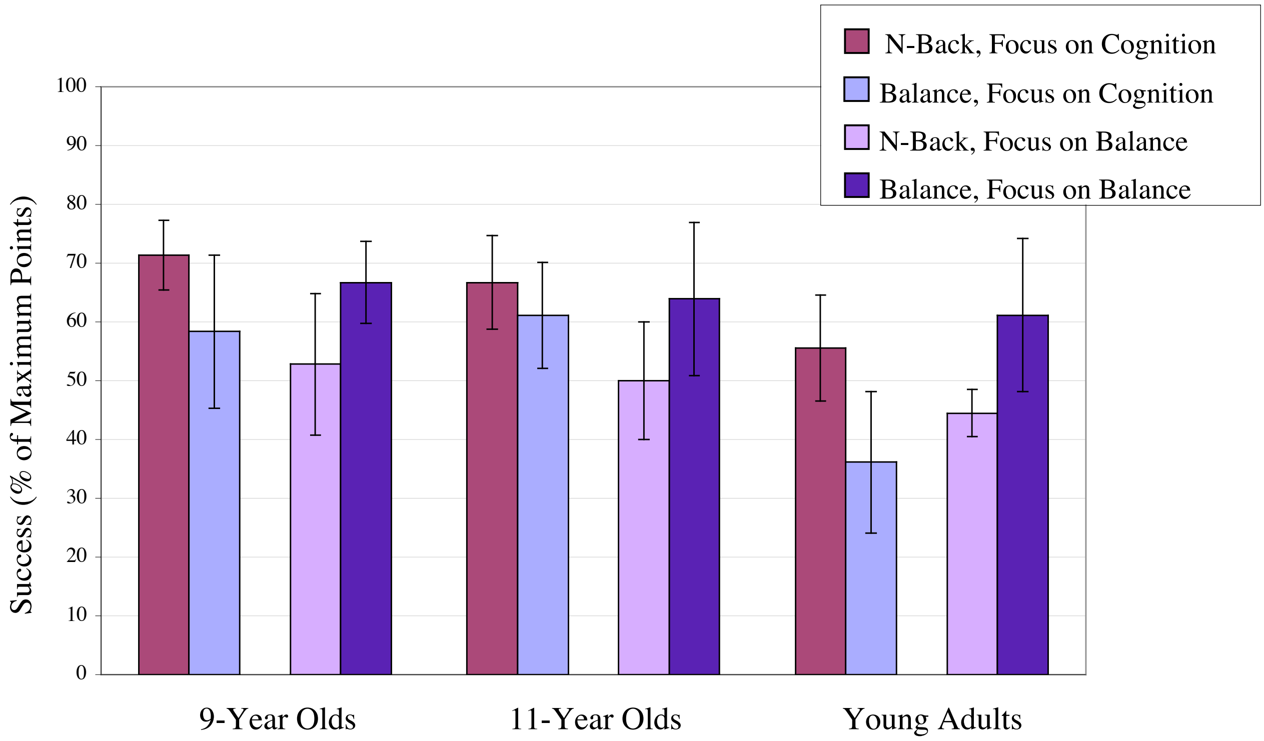

Figure 28. Distribution of Reinforcement Points in the Differential-Emphasis Phase

Note. The percentage of points achieved is depicted in four columns for each age group. The first two columns depict the performance under the instruction to „focus on the cognitive task“, with N-back performance in dark purple and balance performance in light blue. The next two columns depict performance under the instruction to „focus on the balance task“, with N-back performance in light purple and balance performance in dark blue. Error bars depict standard errors of the mean.

In all groups, there was a trend that participants collected more of the potentially achievable points in the task domain on which they should focus on than in the other domain. A mixed-design ANOVA on the proportion of points reached as the dependent variable with task domain (2; cognition vs. balance) and emphasis condition (2; "focus on N-back" vs. "focus on balance") as within-subjects factors and with age group (2) as between-subjects factor was conducted. The effect of task modality was not significant, $F(1,24) = .03$, $MSE = .12$, $p = .875$, $\eta^2 = .001$, and neither was the effect of emphasis-condition, $F(1,24) = .12$, $MSE = .06$, $p = .728$, $\eta^2 = .005$, and none of these effects interacted with the age contrasts. This indicates that participants did not collect more points for one task domain than for the other, or for one emphasis condition as opposed to the other. However, the interaction of the two within-subjects effects of task domain and emphasis condition reached significance, $F(1,24) = 7.56$, $MSE = .07$, $p < .05$, $\eta^2 = .240$. This indicates that participants collected more points for the task domain on which they should focus on than for the other domain. However, the crucial point is that age groups did not differ concerning this interaction effect, ($F(1,24) = .37$, $MSE = .07$, $p = .547$, $\eta^2 = .015$ for children versus adults, and $F(1,24)= .09$, $MSE = .07$, $p = .765$, $\eta^2 = .004$ for 9- versus 11-year olds). Moreover, there were no differences in the overall points reached by children versus adults, $F(1,24) = 3.59$, $MSE = .10$, $p = .707$, $\eta^2 = .130$, or by 9- versus 11-year olds, $F(1,24) = .06$, $MSE = .10$, $p = .803$, $\eta^2 = .003$. This pattern indicates that there are no fundamental differences between children and young adults in their ability to follow differential-emphasis instructions when the proportion of points gained is the dependent variable.

Appendix M presents correlation coefficients for the different DTCs measured in the present study. DTCs did not correlate highly, except DTCs for the balance task, indicating that the prioritization of the balance domain occurred across different dual-task situations.

In sum, the cognitive and balance DTCs of the differential-emphasis phase supported the third hypothesis: Children's DTCs in balance were consistently lower than their DTCs in cognition, even when they were instructed to focus more strongly on the cognitive domain. Young adults' DTCs were comparable in both domains, and they did not show a clear pattern of flexible resource allocation in the their dual-task costs. This shows that children have a harder time than young adults to shift their attention away from the balance domain. On the other hand, the distribution of reinforcement points indicated that all three age groups were able to allocate resources according to instruction. However, the criterion to be reached depended on each individual's previous dual-task performance, and the performance criteria in the children's balancing were therefore in the range of negative DTCs.

5.6 Summary of Main Findings

During training, age differences emerged in the performance of the three component tasks (MOL, N-back, and balance), with young adults showing superior performance to children. Balance difficulty manipulations led to larger COP areas on the moving as compared to the stable platform. The difficulty manipulation of the two cognitive tasks successfully equated task performances in the three groups, and stable performance levels were reached in the single-task trials throughout the dual-task phase of the study.

In line with the prediction, children showed a trade-off pattern in their DTCs, with higher cognitive DTCs than balance DTCs. Young adults' DTCs were comparable across the two domains. At the level of detailed analyses, performance deteriorated in the MOL task in all three groups under dual-task conditions, but the extent of performance deterioration in

MOL did not depend on balance difficulty levels. DTCs were also found in the N-back task in all three age groups, and they additionally depended on balance difficulty: DTCs were higher when people were balancing on the moving as opposed to the stable platform. Balance performance was largely unaffected by concurrent cognitive task in young adults. Children even improved their balance performance under dual-task conditions, resulting in negative DTCs in that task domain.

The second hypothesis postulated that the trade-off pattern in children would be more pronounced when balance difficulty was increased. This pattern was not found in the data, however. Children's dual-task performance trade-offs were very similar on the stable and on the moving platform.

As predicted in the third hypothesis, children were not be able to shift their attention away from the balance task, even when instructed to do so, and continued to show higher DTCs in the cognitive domain as opposed to the balance domain. However, young adults did not show clear evidence for flexible resource allocation, because the differences between their cognitive and sensorimotor DTCs under different emphasis instructions did not always reach significance.