

---

#### 4. METHOD

This study is part of a larger project of the research group “Gender, Resources, and Health in Market- and Family-Work” on the balance of work and family in dual-earner couples with preschool children. Based on the biopsychological model of stress as proposed by Frankenhaeuser (e.g. Frankenhaeuser, 1994) the emphasis of this project is on internal and external factors contributing to the combination of work and family in employed parents and these factors’ relationship with measures of well-being at different levels of functioning. Within this broader framework the main emphasis of the present study is on factors that contribute to the successful management of work- and family-related goals in employed parents with preschool children and its relationship to psychological and physiological measures of well-being. Because it can be assumed that the partner plays a central role in the successful combination of work and family goals, both partners’ perspectives are investigated. Due to the fact that research on intergenerational relations indicates that grandparental participation in childcare is a key resource for the successful combination of work and family in employed parents, special attention is given to grandparental childcare as an additional source of support within the present sample.

The empirical investigation comprises three measurement contacts. At the first measurement contact, employed parents were asked to complete a set of questionnaires to assess the main predictors, namely differences in the quality of interpersonal goal relations, individual time-characteristics of work and family goals, and access to grandparental participation in childcare. At the second measurement contact, employed parents were asked to participate in a one week intensive study of daily life using time-sampling methods. The purpose of this part was to investigate whether the quality of interpersonal goal relations within the partnership and access to grandparental support in the form of childcare influence everyday goal pursuit in husbands and wives and whether differences in goal-relevant actions would be reflected by current affect quality as well as cortisol secretion. At the third measurement contact, study participants were asked to complete questionnaires of outcome measures covering overall goal progress and different indicators of subjective well-being.

In the following section I will present sample characteristics, employed measures, and the procedure of this investigation. A general overview of the study is depicted in Figure 3.

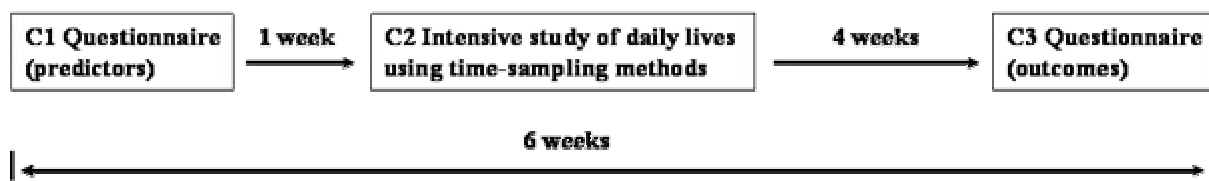


Figure 3. *Study overview*

## 4.1 Sample

### 4.1.1 Recruitment

Employed parents were recruited by advertisements in local newspapers, information leaflets at pediatrician offices, and through public and private organizations in Berlin. Requirements for participation were (1) that both partners agreed to participate in all parts of the study, (2) that both partners worked a minimum of 20 hours per week, (3) that they had at least one preschool child, and (4) that both partners had at least Abitur (13 years of education). These factors were checked in a first telephone interview.

If requirements were met, participants were screened for factors that interfere with the interpretation of cortisol or participation in the time-sampling part of this study. On the basis of this screening several participants had to be excluded. Criteria for exclusion were (1) a job that required them to be away from the family during the whole week, (2) illnesses and medication that have an influence on the hormonal system such as diabetes or thyroid diseases, (3) pregnancy or breastfeeding.

Instead of a monetary reimbursement, employed parents were offered feedback about their time-management and information about factors that seem to contribute to a successful combination of work and family in the present sample. Additionally, all participating couples took part in a lottery. The winner could spend a wellness weekend with their family.

### 4.1.2 Participants

The sample is composed of 53 dual-earner couples. Employed parents had an average age of 37 years ( $SD=4.9$ ), and number of children ranged from one to four with a mean of 1.7. Childcare by others ranged from 5 to 45 hours per week ( $M=35.6$  hours). The majority of the sample was married (83 %), 17 percent were cohabiting. Ninety percent held a university degree, 10 percent had received other vocational training. Average time on the job ranged

from 14 to 71 hours per week ( $M= 41.0$  hours) with women spending an average of 8 hours less in employment than men. Half of the sample (48 %) had an income of more than 2000 Euro per month, which was significantly higher in men as compared to women. Detailed descriptions of the sociodemographic characteristics are presented in Table 4.

Table 4. *Sociodemographic Characteristics: Employed Parents*

	Female (N=53)	Male (N=53)	All (N=106)
<b>Age (in years)</b>			
Range	25.6 – 44.7	26.8 – 48.9	25.7 - 48.9
Mean	36.1	37.4	36.8
SD	4.6	5.2	4.9
<b>Marital status</b>			
Married	44 (83.0 %)	44 (83.0 %)	88 (83.0 %)
Cohabiting	9 (17.0 %)	9 (17.0 %)	18 (17.0 %)
<b>Number of children</b>			
Range	1 – 4	1 – 4	1 – 4
Mean	1.7	1.7	1.7
SD	0.64	0.64	0.64
<b>Age of children (in years)</b>			
Range	0.0 – 9.5	0.0 – 9.5	0.0 – 9.5
Mean	4.8	4.8	4.8
SD	2.1	2.1	2.1
<b>Childcare by others (hours per week)</b>			
Range	5 – 45	5 – 45	5 – 45
Mean	35.6	35.6	35.6
SD	7.1	7.1	7.1
<b>Access to grandparental childcare</b>			
Yes	30 (56.6 %)	32 (60.4 %)	62 (58.5 %)
No	22 (41.5 %)	21 (39.6 %)	43 (40.6 %)
<b>Education</b>			
University / University of applied sciences	48 (90.6 %)	47 (90.4 %)	95 (90.4 %)
Other professional training	5 (9.4 %)	5 (9.6 %)	10 (9.5 %)

*(table continues)*

Table 4. (continued)

	Female (N=53)	Male (N=53)	All (N=106)
<b>Income (per month)</b>			
Less than 1000 Euro	8 (15.4 %)	3 (5.7 %)	11 (10.5 %)
1000 to 1500 Euro	21 (40.4 %)	3 (5.7 %)	24 (22.9 %)
1500 to 2000 Euro	6 (11.5 %)	16 (30.2 %)	22 (21.0 %)
More than 2000 Euro	17 (32.7 %)	31 (58.5 %)	48 (45.7 %)
<b>Time on the job (hours per week)</b>			
Range	14 – 70	15 - 71	14 - 71
Mean	36.8	45.2	41.0
SD	11.51	12.6	12.7

### 4.1.3 Sample Selectivity

Given that this study is based on a sample of highly educated dual-earner couples with small children, two questions that address different aspects of sample selectivity will be addressed below. The first refers to the choice of recruitment strategies and the second to a description of actual and desired employment patterns among German couples with preschool children.

In order to recruit a variety of couples out of the target population, multiple recruitment strategies were employed (see above). While it is impossible to estimate what percentage of the couples who knew about the study were willing to donate their time to research, records were kept on the number of couples who (1) contacted us but did not participate due to time constraints, who (2) were excluded based on the initial screening, and who (3) dropped out during time in study. Out of a total of 133 couples that called in, 42 ended up not participating mainly because of time-constraints for one partner. 23 couples had to be excluded based on the screening information and 15 dropped out during the study.

I will next turn to the question of actual and desired employment patterns among German couples with children under the age of 6 years. According to the Organization for Economic Cooperation and Development (OECD; 2001), which provides information about actual and desired employment patterns in German couples, dual-earner couples where both partners are employed more than 16 h/week represent 15.7 percent of the German population of couples with children under the age of six. The majority of parents with children of this age-group still adhere to a pattern where the man holds the role of the breadwinner and the woman does not participate in the labor market. Despite these current patterns, one third of this population would prefer an employment pattern where both partners are employed more

than 16 h per week. An additional 42.9 percent report desiring a situation where the husband is employed full time and the wife part time.

According to the same survey (OECD, 2001), the labor force participation of mothers with children below the age of six differs by education. While 62.4 percent of the mothers holding a university degree were employed, only 28.7 percent of mothers without professional training were employed. Hence, women with high levels of education like the participants of the present study are twice as likely to work for pay as compared to women with low levels of education, at least when children are still small. Another important factor that seems to be positively associated with the participation of mothers in the labor force is the availability of public childcare. Especially daycare facilities for small children that are open not only in the morning but also in the afternoon prove to be an important factor for mothers' participation in the labor force (Buechel & Spiess, 2002). Because Berlin is the state in Germany which has by far the highest supply of public childcare, participants of the present study are privileged in this respect compared to employed parents living in other areas of Germany (Buechel & Spiess, 2002).

Hence, despite the fact that the present sample is select in terms of level of education and women's labor-force participation, part of which might be attributable to the high availability of public childcare specific to Berlin, participants still practice a model of combining work and family that seems to be considered desirable by a substantial proportion of their age-group.

## **4.2 First Measurement Contact**

### **4.2.1 Procedure**

At the beginning of the study, employed parents were contacted by phone and asked which of the upcoming weeks suited them to participate in the time-sampling study. Two weeks prior to this individually scheduled week, both partners were sent an e-mail providing them with a link to the first package of questionnaires on the Internet. Within that e-mail they were also told that they could complete the questionnaires at different times but answer the whole set within one week. Study participants were asked to complete the questionnaires alone. The Internet questionnaires took approximately 3 hours to complete.

## 4.2.2 Measures

In the following section the measures of the first set of questionnaires are introduced. Because this study is part of a bigger research project, a variety of instruments has been employed. However, only the ones relevant for this study will be described. Table 5 gives an overview of the central measures of the first measurement contact.

Table 5. *Overview of CI Measures*

Construct	Instrument	Source
Predictors:		
Personal goals	Personal Project Analysis	Modification of Little, 1983
Interpersonal goal relations (among the 4 work and family goals of both partners)	Interpersonal goal questionnaire	Newly developed
Time-intensity of work and family goals	1 Item	Newly developed
Flexibility in the pursuit of work and family goals	1 Item	Newly developed
Grandparental involvement in childcare	1 Item	Newly developed
Outcomes:		
Emotional well-being	Multidimensional Affect Scale	Steyer, Schwenkmezger, Notz, & Eid, 1997
Psychological well-being	Ryff-Scales	Ryff, 1989 in the German translation of the LIP-project of the Max-Planck-Institute of Human Development, Berlin

### 4.2.2.1 Personal Goals

In the assessment of personal goals a two-step procedure was chosen. Employed parents were asked to list all the goals they currently were thinking about, planning for or carrying out. In order to restrict the hierarchical level of the reported goals and to elicit goals that can be assumed to be related to the everyday behavior that is of particular interest within this study, the original instruction by Little (1983) was further specified. According to this specification, study participants were asked to report those goals (1) that they actively planned to pursue within the next weeks, (2) that were already important to them at the moment of data collection, and (3) that were related to their everyday activities. Personal goals were categorized by two independent raters along a modified coding scheme developed by Riediger (2001). The system specified 17 different categories. Interrater agreement was high (Cohen's Kappa  $k = .92$ ). Figure 4 shows the 10 content categories that personal goals most often referred to in the present sample. Because participants have named different

numbers of goals (range: 2 – 11), proportions of goals per content category are presented. In line with previous studies, work and family were the categories that were most often represented in the personal goals of employed parents with preschool children.

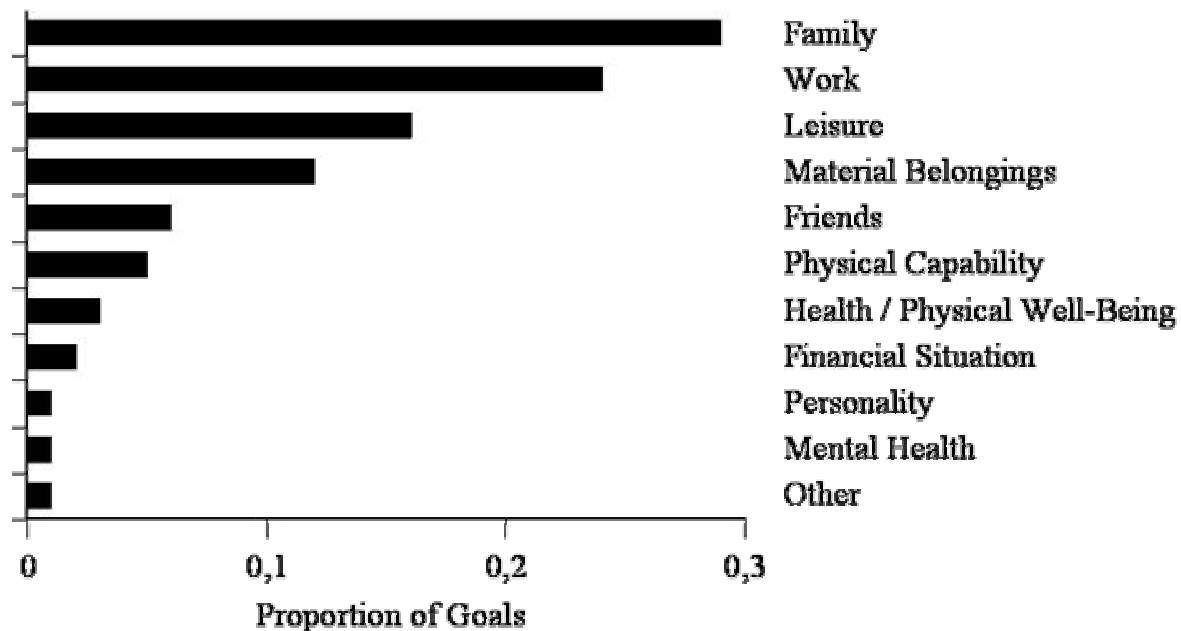


Figure 4. *Content categories of personal goals of employed parents*

Due to a focus on work and family goals, employed parents were secondly asked about their most important goals with respect to four content categories, namely (1) one partnership-related goal, (2) one child-related goal, (3) one goal that was related to their professional development and (4) one goal that referred to current work projects and tasks. These content categories were obtained from a secondary analysis of the personal goals that 20 randomly chosen parents (10 women and 10 men) with small children had named in a study by Salmela-Aro and colleagues (Salmela-Aro, Nurmi, Saisto, & Halmesmäki, 2000).

#### 4.2.2.2 Interpersonal Goal Relations

*Assessment procedure.* For the assessment of the quality of interpersonal goal relations (based on two work- and two family-related goals) a modified version of the Personal Projects Matrix (Little, 1983) was used. In that matrix the row headings represented the content categories of study participants' own work- and family-related goals, whereas the column headings depicted the content categories of the presumed goals of the respective partner.

Because I was interested in two components of interpersonal goal convergence (time-based and strategy-based convergence) and two components of interpersonal goal conflict (time-based and strategy-based conflict) four matrices were presented, each of them headed by a different question. The first question addressed interpersonal goal convergence as related to resource expansion and read: How often does it happen that you have more time for the pursuit of goal X because your partner is working on his goal Y? The second question addressed interpersonal goal convergence as related to action facilitation and read: How often does it happen that your partner's pursuit of goal Y simultaneously furthers your own goal X? The third question addressed interpersonal goal conflict as related to resource limitation and read: How often does it happen that you do not have enough time for the pursuit of goal X because your partner is working on his goal Y? And the fourth question addressed interpersonal goal conflict as related to action hindrance and read: How often does it happen that your partner's pursuit of goal Y simultaneously hinders your own goal X? Study participants filled in the cells of each matrix using a scale ranging from never (1) to very often (5). In case any two goals were independent of each other, study participants were instructed to indicate "1" for "never".

A subsample of 15 couples also completed a simplified version of this questionnaire. The assessment procedure paralleled the above described full version with the exception that the two work goals and the two family goals were subsumed under one work goal category and one family goal category resulting in two by two instead of four by four matrices.

*Item characteristics.* Overall Item characteristics of the interpersonal goal conflict and convergence questionnaire were satisfactory. Table 1 in Appendix A contains detailed descriptions for each of the 64 items. Item difficulties covered the whole range ( $P_{con} = .06$  to  $.98$ ) reflecting a high variability in responses. Apart from four exceptions, the item discriminabilities were in an acceptable range ( $r_{con} = .15$  to  $.78$ ), warranting their use for scale aggregation.

*Treatment of missing values and outliers.* In case of missing values I applied the following procedure: If at least 50% of the subscale items were available, missing values were estimated using regressions by gender and the response to the other items. If more than 50% of the subscale items were missing, the subscale score was set to missing. If participants with more than 50% missing values on a subscale belonged to the sub sample who answered a short form of the interpersonal goal relations questionnaire, the subscale of the long version was substituted by the respective subscale of the short version. This procedure seemed justifiable because subscale correlations between the short and long versions of the



interpersonal goal relations questionnaire were reliable ( $r = .49$  to  $.87$ ). The number of missing values ranged from 0 to 15 with a mean of 1.48 missing values per sub scale. Following recommendations by Tabachnick and Fidell (1996), values that were more than 1.5 box lengths larger than the 75<sup>th</sup> percentile or 1.5 box lengths smaller than the 25<sup>th</sup> percentile were treated as outliers and adjusted to the closest non-outlying value in the data set.

*Subscale aggregation.* Prior to subscale aggregation, independent goals were excluded from the data. This was done by setting those values to missing where study participants had indicated that a particular goal pair led neither to time-based goal convergence nor to time-based goal conflict. The same procedure was taken if neither strategy-based goal convergence nor strategy-based goal conflict applied.

Averaging the same item across all goal pairs yielded the two interpersonal goal conflict (time-based and strategy-based goal conflict) and two interpersonal convergence (time-based and strategy-based goal convergence) subscales. Because examination of subscale distributions revealed departures from normality in three of the four subscales, square root transformations were applied to the subscales (1) strategy-based goal conflict, (2) time-based convergence, and (3) strategy-based goal convergence. This yielded satisfactory symmetric distributions as indicated by absolute ratios of skewness and kurtosis to their respective standard errors being smaller than two (Tabachnick & Fidell, 1996, see Table 2 in Appendix A). Cronbach's Alphas showed satisfactory internal consistencies of the four subscales (see Table 6).

Table 6. *Cronbach's Alpha for the Conflict and Convergence Subscales*

Subscale	N Items	Coefficient
Time-based conflict	16	.86
Strategy-based conflict	16	.91
Time-based convergence	16	.82
Strategy-based convergence	16	.86

Because correlations between the subscales without estimated missings, with estimated missings, and with estimated missings plus transformations were very high (see Table 3 in Appendix A), analyses of the questionnaire structure reported below refer to the transformed subscales with estimated missings. Univariate outliers were adjusted to the closest non-outlying value in the data distribution based on recommendations by Tabachnick and Fidell (1999). According to this procedure three cases were adjusted in the time-based

conflict subscale, two cases were adjusted in the time-based and strategy-based convergence subscales and one case was adjusted in the strategy-based conflict subscale.

*Subscale correlations.* As expected, the two conflict and the two convergence subscales were reliably positively correlated. The correlations between conflict and convergence subscales were negative but not as high as they would be if they represented opposite poles on the same dimension.

Table 7. *Correlations of the Interpersonal Conflict and Convergence Subscales*

	Conflict		Convergence	
	I	II	I	II
Conflict				
I Time-based	1.00			
II Strategy-based	.76 **	1.00		
Convergence				
I Time-based	-.14	-.02	1.00	
II Strategy-based	-.22 *	-.46 **	.40 **	1.00

\*\*  $p < .01$ ; \*  $p < .05$

*Exploratory factor analysis.* Principal components extraction with varimax rotation was employed to further explore the structure of the interpersonal goal relations questionnaire (see Table 8). The analyses yielded two components with Eigenvalues greater than one. Component loadings revealed that both conflict subscales loaded on the first component and both convergence subscales loaded on the second component. The only convergence subscale that also loaded negatively on the first component was strategy-based convergence.

Table 8. *Principal Component Analyses with Varimax Rotation of the Conflict and Convergence Subscales*

Subscales	Component 1 (Conflict)	Component 2 (Convergence)	Communality
Time-based conflict	<b>.90</b>	-.06	.81
Strategy-based conflict	<b>.94</b>	-.13	.91
Time-based convergence	.09	<b>.90</b>	.81
Strategy-based convergence	-.40	<b>.74</b>	.71
% explained variance	46.40	34.33	
Cumulative % explained variance	46.40	80.73	

In order to test whether the observed two-factor structure of the instrument was not only present in the entire set of goals but also with respect to specific sets of goals, separate principal component analyses were performed on each of the 16 pairs of goals. These analyses uniformly showed that the two conflict items loaded positively on one component and the two convergence items loaded positively on the other component, thus replicating the above presented structure. The two factors together accounted for 73.43 to 89.17 percent of the variance.

Based on the described correlational and factor-analytic results, the four subscales were collapsed into one conflict and one convergence scale. These two scales were only moderately negatively correlated ( $r = -.33^{**}$ ), supporting the notion that they do not present opposite poles on a single dimension.

*Ruling out a methodological concern.* From a methodological point of view it seemed important to provide evidence that the newly developed interpersonal goal relations questionnaire does not represent a new measure of relationship satisfaction or social support. I therefore investigated its association with the relationship assessment scale (Sander & Boecker, 1993) and two measures of social support, namely social support by others (Cohen, Mermelstein, Kamarck, & Hoberman, 1985) and goal support on side of partner (newly developed). Pearson's correlations between interpersonal goal relations and relationship satisfaction revealed both, interpersonal goal conflict ( $r = -.17$  ns) and interpersonal goal convergence ( $r = .05$  ns), to be independent of relationship satisfaction. Additionally, no significant relationship was found between interpersonal goal convergence and social support by others ( $r = .06$  ns) and goal support on side of partner ( $r = .10$  ns). Similarly, interpersonal goal conflict correlated neither with social support by others ( $r = -.06$  ns) nor with goal support by partner ( $r = -.06$  ns). These results support the notion that the newly developed questionnaire is not identical with preexisting measures of relationship satisfaction or social support. Correlations with additional personality and self-regulatory measures are displayed in Table 4 in Appendix A.

#### **4.2.2.3 Antecedents of the Quality of Interpersonal Goal Relations**

*Time-intensity of work and family goals.* The time-intensity of work and family goals was assessed separately for each goal from the four categories. Study participants were asked to indicate how much time (in hours and minutes) per week they would need to spend working on each particular goal if they wanted to achieve it (see Table 5 in Appendix A for distributions).

*Flexibility in the pursuit of work and family goals.* For the assessment of differences in the temporal flexibility of goal pursuit, a matrix was developed based on which the construct of interest could be calculated. The development of the matrix was informed by assessment procedures that have been used for the assessment of the flexibility of household tasks (Barnett & Shen, 1997; Resch, 1999; Valax, 1998). Study participants were presented a weekly schedule with days in rows and hours in columns. They were asked to check all cells where they could work on their goals given an average week. For each of the two work- and two family-related goals, separate matrices were completed. Flexibility scores were obtained by a division of the time-intensity of a given goal by the amount of time that could be devoted to goal pursuit. Flexibility scores were computed separately for each of the four goals and then averaged into a single score (see Table 5 in Appendix A for distributions).

#### **4.2.2.4 Grandparental Involvement in Childcare**

Grandparental involvement in childcare was assessed by asking employed parents if they had access to a grandparent to whom they could turn for childcare. If so, they were further asked to specify how this person was related to them. Additionally, sociodemographic characteristics for all living grandparents were collected.

#### **4.2.2.5 Subjective Well-Being**

At C1 employed parents completed two instruments assessing emotional and cognitive aspects of well-being, namely (1) emotional well-being and (2) psychological well-being.

*Emotional well-being.* The multidimensional affect scale (Steyer et al., 1997) was used to assess the experience of different affect qualities prior to participation in the study. The multidimensional affect rating distinguishes between positive mood, ease, and alertness as positive affect dimensions and negative mood, restlessness, and fatigue as negative affect dimensions. Each of these affect dimensions is represented by four adjectives. Study participants were asked to indicate on a scale ranging from very seldom (1) to not at all (5) how often they had experienced each affect quality during the three to four month before the initial assessment.

*Psychological well-being.* The Ryff-Scales assess six dimensions of psychological well-being, each of them being represented by nine items (Ryff, 1989). The subscales are: (1) self-acceptance (e.g. “I like most parts of my personality.”), (2) personal growth (e.g. “I think it is important to have new experiences that challenge how I think about myself and the

world.”), (3) purpose in life (e.g. “Some people wander aimlessly through life; I am not one of them.”), (4) environmental mastery (e.g. “In general, I feel I am in charge of the situation in which I live.”), (5) autonomy (e.g. “I judge myself by what I think is important, not by the values of what others think is important.”), and (6) positive relations with others (e.g. “People would describe me as a giving person, willing to share my time with others.”). Response options ranged from not at all (1) to very much (5). In this study the German translation by the LIP-group at the Max-Planck-Institute of Human Development, Berlin was used.

### **4.3 Second Measurement Contact**

#### **4.3.1 Procedure**

The second measurement contact began approximately one week after completion of the C1 package of questionnaires on the Internet. On an individually scheduled evening, one of three trained research assistants went to the family home and introduced both partners to the use of the study device. The instruction session took about 1.5 hours per couple. Within this session the research assistant explained the details of the measurement procedure and introduced study participants to the use of the pocket computers (Psion series 3a) and the Salivette sampling device (Sarstedt, Rommelsdorf, Germany). The research assistant went through each questionnaire from the time-sampling phase in great detail and clarified any questions of the study participants. Both partners also received a paper summary containing all instructions and a cell-phone number for further questions during time in study.

After having been familiarized with the study materials, study participants completed a sample questionnaire themselves. At the end of the session, research assistants set the alarms for the time-sampling phase, which started on the morning following the introductory session.

The time-sampling phase covered six consecutive days (always including four week days and both weekend days). Each day started with a morning questionnaire. The morning questionnaire was completed right after waking up and included affect ratings, reports of activities participants had been engaged in after the last report on the evening before, as well as questions concerning their goal-relevance. Parallel to questionnaire completion, participants were asked to provide a saliva sample for estimation of cortisol. During the day (between 9.30 a.m. and 9.30 p.m) both partners simultaneously received five different beeps separated by intervals of approximately three hours. After each alarm, study participants filled in the day questionnaire containing affect ratings, their activities after the last beep, and

---

questions concerning the goal-relevance of current and past activities. Parallel to questionnaire completion they were again asked to provide saliva samples. Participants were asked to answer the beeped questionnaire as soon as possible after the alarm. Both the morning- and the day-questionnaires took five to seven minutes to complete. At the end of the time-sampling phase, a research assistant collected the study materials at the family home and both partners were asked to complete a short reactivity questionnaire.

### **4.3.2 Measures**

#### **4.3.2.1 Morning Questionnaire**

The morning questionnaire consisted of the following parts: First, study participants answered six questions from the Multidimensional Affect Scale (Steyer et al., 1997) for assessment of current affect quality. Items were chosen to represent each of the three affect dimensions with one positive and one negative adjective. Affect questions were presented at the beginning of the questionnaire in order to avoid being affected by completion of the other parts of the questionnaire.

In the second part, employed parents were asked to list the activities they had been engaged in the previous evening after the last beep. They were asked to chronologically report their activities in 15minute time-intervals on a preprinted table using abbreviations from a category system. The system is summarized in Table 9. Because of its complexity, it was explained and trained on examples during the introductory session. This part was the only part which was done on paper.

The third part of the morning questionnaire addressed the goal relevance of these reported everyday life activities. Study participants were asked for each of their four work- and family-related goals separately, which of the listed activities had been goal-relevant and if so, in what way ranging from “activity very much hindered that goal” (-3) to “activity very much furthered that goal” (+3).

At the end, study participants were asked to report the number of the saliva sample they had used parallel to questionnaire completion. Additionally, there was room for comments, e.g., on special circumstances.

Table 9. *Categorical System*

Categories	Subcategories (Abbreviations)
(1) Market work	Commuting to work (E1), working with other colleagues on tasks (E2), working alone on tasks (E3), further training (E4), working on tasks with customers (E5), paper work (E6), break (E7), private conversation (E8), and commuting at work (E9)
(2) Leisure time	TV/radio (F1), hobbies/cultural activities (F2), physical activities (F3), reading/computer (F4), spirituality (F5), volunteer work (F6), social contacts (F7), and commuting (F8)
(3) Household work	Meal preparation and cleaning up after meals (H1), shopping (H2), house-/car-/garden-maintenance (H3), paper work and organizational tasks (H4), cleaning the house (H5), support of family members in need of care (H6), doing laundry (H7), and commuting (H8)
(4) Childcare	Supervising children (K1), activities with children (K2), health care of children (K3), taking care that children are getting washed and dressed (K4), and commuting (K5)
(5) Personal things	Intimacy with partner (P1), relaxation (P2), eating (P3), medical care (4), self care (P5), and other (P6)

#### 4.3.2.2 Day Questionnaire

The items from the day questionnaire that are relevant for the present study are parallel to the morning questionnaire. The first block consisted of the same set of affect items from the Multidimensional Affect Scale (Steyer et al., 1997). The second part addressed the different activities participants had been engaged in during the preceding hours after the last beep along the above described coding system (see 4.3.2.1). Following this they were asked about the goal-relevance of these everyday activities with respect to the four previously listed work- and family-goals (see 4.3.2.1). And at the end of the questionnaire, study participants were again asked to provide the number of the used saliva sample and asked about special circumstances.

#### 4.3.2.3 Scale Aggregation

*Affect quality in everyday life.* The morning and the day questionnaires contained two items (1 positive and 1 negative) for each of the three subscales of the Multidimensional Affect Scale (Steyer et al., 1997). Prior to subscale aggregation the three negative items were recoded. The three subscales of positive-negative mood, ease-restlessness, and alertness-fatigue were aggregated by averaging responses to the respective items. Descriptives of these three scales are displayed in Table 6 in Appendix A.

*Goal relevance of everyday activities.* Participants had been asked to chronologically report the goal-relevance of their everyday activities. For each activity, they were asked to specify if any particular activity had been relevant for each of the two work and two family goals and if so, whether it had furthered or hindered them in achieving that particular goal using a scale ranging from -3 (very much hindered) to +3 (very much furthered). Goal relevance of everyday activities was aggregated by calculating the mean goal-relevance of all goal-relevant activities during a particular interval. Scale descriptives are provided in Table 6 in Appendix A.

*Physiological arousal.* Participants had provided saliva samples on every measurement point during the time-sampling period. Two measures of cortisol were of particular interests within the framework of this study. The first measure addresses moment-to-moment fluctuations in cortisol. Because cortisol undergoes a diurnal cycle, person-specific difference scores between single cortisol values and the average of all cortisol values taken at the same time of the day over the 6-day period were computed.

The second measure is the so called “area under the curve” (Pruessner, Kirschbaum, Meinlschmid, & Hellhammer, 2003). This construct reflects the overall output of cortisol during an observation period. It was computed for each day in study using the trapezoid formula developed by Pruessner and colleagues (Pruessner et al., 2003). Cortisol is measured in nmol/l.

### **4.3.3 Reactivity**

After completion of the time-sampling phase, participants were asked about their experience with the intensive assessment procedure. Hence, a feedback-questionnaire was administered that covered different aspects that have important implications for evaluating the validity of the obtained data, namely acceptance of the time-sampling methodology, accuracy of answers, and reactivity. Employed questions and participants’ responses are summarized in Table 10. In addition to that we also asked study participants whether they alternate in the engagement in work and family and, if so, about the according time-spans. I will first turn to a description of the acceptance of the procedure.

Overall, results on the different acceptance questions are in accordance with observations by other research groups using time-sampling methodologies in the family setting (e.g. Perez, Schoebi, & Wilhelm, 2000). They show that the intensive assessment procedure did not disturb participants in their daily lives. The majority of participants indicated that the study device was handy and that they did not experience negative reactions



by others when they filled out the questionnaires. Two thirds of the sample said that the procedure did not interrupt their daily routines, while a third admitted that they did find it disruptive.

Table 10. *Feedback on the Experience during the Time-Sampling Phase*

	<b>Percent endorsement</b>
<b>Acceptance</b>	
Ease of use	93.7
Unpleasant reactions by others	4.3
Disruption of daily routines	29.5
<b>Accuracy</b>	
Difficulty with questionnaire completion	1.1
Comprehensibility of questions	96.8
Typicality of days in study	85.4
Difficulty in classification of activities	3.3
Difficulty with length of activity intervals	8.5
Minutes delay of questionnaire completion (in % beeped measurement points)	
Mean (SD) min	30.0 (56.9)
< 15 min	63.7
15 – 30 min	8.6
30 – 60 min	9.4
60 – 120 min	13.2
> 120 min	5.0
Missing data (in % beeped measurement points)	3.5
<b>Reactivity</b>	
Behavior change	5.3
Greater frequency of talking about goals with partner	29.5
Avoidance of reporting conflicting goals in interpersonal goal questionnaire	0.0

Concerns with respect to the accuracy of reports cover aspects such as the comprehension of the content of questions, the typicality of the time-sampling period compared to any other week, and the delay in questionnaire completion as well as missing data. The majority of the sample reported having had trouble neither in understanding the questions nor in categorizing the activities they were engaged in on the basis of 15 minute intervals. Results also reveal that we were successful in sampling a week that participants considered to be typical of their daily lives. The overall commitment to the procedure was

very high as indicated by a very low percentage (3.5 %) of failures to report. Two thirds of the sample answered the questionnaires within 15 minutes after the beep. The remaining third delayed their responses by up to two hours. Within the context of this study, delayed answers were treated as indicators of low commitment but remained in the data because no predictions were made with respect to the recency of a given event.

A very crucial aspect in any time-sampling study is the question as to what extent participants change their behavior in response to the topic under investigation. In the present study not a single participant indicated having avoided to report conflicting work and family goals in the beginning of the study. Only five percent of the sample said that they think their behavior had changed in response to the questions asked. And a third of the sample reported having talked more often than usual about their personal goals with their partner.

A final concern addresses the question of how extensively partners alternate in their involvement in work and family. We therefore asked our study participants whether they rotate on a regular basis and, if so, how much time these cycles span. Of the 33 couples who indicated that they alternated on a regular basis, only one couple said that they rotated per week and no couple reported alternations referring to longer durations than that. Hence, it seems justified to assume that we did manage to capture an average week that represents the daily lives of the sample under study.

#### **4.4 Third Measurement Contact**

##### **4.4.1 Procedure**

The third measurement contact took place one month after completion of the intensive study of daily lives. As was the case at C1, employed parents were sent an e-mail with a link to the outcome questionnaire at the homepage of the research group and asked to complete the questionnaire alone within one week. Completion of the second questionnaire took about 45 Minutes.

##### **4.4.2 Measures**

I will now review the outcome measures. The employed instruments are summarized in Table 11.

Table 11. *Overview of C3 Measures*

Construct	Instrument	Source
<b>Outcomes:</b>		
Emotional well-being	Multidimensional Affect Scale	Steyer et al., 1997
Psychological well-being	Ryff-Scales	Ryff, 1989 in the German translation of the LIP-project of the Max-Planck-Institute of Human Development, Berlin
Goal progress	Single item	Riediger, 2001
Goal-specific satisfaction	Single item	Riediger, 2001

#### 4.4.2.1 Subjective Well-Being

At C3 employed parents completed two instruments assessing emotional and cognitive aspects of well-being, namely (1) emotional well-being and (2) psychological well-being.

*Emotional well-being.* The long version of the Multidimensional Affect Scale (Steyer et al., 1997) was used to assess the experience of different affect qualities during time in study. Descriptives are displayed in Table 7 in Appendix A.

*Psychological well-being.* The full version of the Ryff-Scales was administered to assess six dimensions of psychological well-being (Ryff, 1989). Descriptives are displayed in Table 7 in Appendix A.

#### 4.4.2.2 Goal Progress

Employed parents were asked to indicate on a 7-point scale whether they had moved toward or away from each of their four work- and family-related goals since the first measurement contact. Response options ranged from “moved very much away” (1) from that goal to “moved very much toward” (7) it (Riediger, 2001). Descriptives are displayed in Table 7 in Appendix A.

#### 4.4.2.3 Goal-Specific Satisfaction

Study participants were asked to report their satisfaction with the development on each of their work and family goals along a scale ranging from “very dissatisfied” (1) to “very satisfied” (7). This item was developed by Riediger (2001). Averaging answers across the four responses yielded the individual goal-satisfaction scores. Descriptives are displayed in Table 7 in Appendix A.

#### **4.5 Advantages and Caveats of Assessments via Internet**

Because most participants had access to computers both at work and at home, we chose to conduct the questionnaires of the first and third measurement point via Internet. This assessment procedure has the advantage that data can be downloaded without delay and that one does not have to worry about mistakes that can occur in transferring paper-based answers into a computerized format.

However, a disadvantage of Internet based assessment procedures lies in the low controllability of missing answers. Within the framework of this study, 13 participants (12 % of the sample) did not provide enough information on the central constructs of this study to be included in the analyses. Participants were excluded from analyses if more than 50 percent of the items behind one of the central constructs were missing.

Because the main interest of this study was in investigating the interrelationship among predictors and criteria in husbands and wives, we always had to exclude the whole couple whenever one partner's answers were missing. This resulted in a reduction of the number of couples by 11 (21 %). Hence, analyses presented in the result section are based on a total of 84 individuals in 42 couples who provided complete data sets.

#### **4.6 Data Analytic Strategies**

Data obtained in research on couples are non-independent in nature. Non-independence of data undermines the statistical assumption that observations are independent replications. While mean differences and regression coefficients are not affected by non-independence in data, standard errors, p-values, confidence intervals, and effect-sizes are invalid. The biasing effect of non-independence can make significance testing either too liberal or too conservative (Snijders & Bosker, 1999).

Because theory and design of the present study make non-independence of data very likely, it is important to use statistical models that take care of the assumed dependencies between observations. Multilevel analysis is a methodology for the analysis of data with complex patterns of variability. It is especially useful in modeling data with a hierarchically nested structure.

The present study assumes a three level structure in the data. The first level concerns repeated assessments in daily life, the second level individuals, and the third level couples. Using multilevel analysis it is then possible to estimate parameters within a certain level of nesting and to also examine relations between the different levels.

Within the present study the Actor-Partner-Interdependence Model (APIM; Kashy & Kenny, 2000) served as a framework to investigate the assumed mutual influences between individuals in partnerships (see section 2.2.3.). The assumption underlying this model is that a person's independent variable score affects not only his or her own dependent variable score (actor effect) but also that of the respective partner (partner effect) and that individual scores also depend on differences in couple constellations (compositional effects). Hierarchical Linear Modeling allows modeling of actor- and partner- as well as compositional effects simultaneously (HLM; Bryk & Raudenbush, 1992; Campbell & Kashy, 2002).

The simplest version of the APIM assumes that actor and partner are interchangeable. Because the present study consists of employed parents that differ at minimum with respect to their sex, I paid particular attention to the influence of gender differences. Hence, gender was always included in the analyses. Additionally, gender differences in the structure of relationships were investigated. In case gender differences were detected, separate coefficients were estimated for husbands and wives (Raudenbush, Brennan, & Barnett, 1995).

I will now turn to a description of the procedure that is employed for hypotheses testing: First, intraclass correlations are estimated to specify the amount of variability at each level using fully unconditional models (Snijders & Bosker, 1999). Then, conditional models are estimated including all theoretically relevant predictors. Predictors at the first level usually refer to repeated observations in daily life and are centered around each person's mean on that variable. Predictors at the second level reflect individual difference characteristics and are centered around the overall mean of all study participants. Predictors at the third level refer to couple characteristics and are centered around the mean of all participating couples. Centering is insofar important as it sets the intercept to a meaningful value facilitating the interpretation of results (Campbell & Kashy, 2002; Hox, 2002; Schwartz & Stone, 1998). Within the present study none of the variables is standardized. Hence results have to be interpreted in reference to the original scaling of predictors and criteria. Assumptions underlying multi-level modeling are checked by means of screening the scatter plots of residuals against their predicted values (Hox, 2002; Tabachnick & Fidell, 1996). Then different covariance structures are modeled in order to compare coefficients and standard errors between models based on unstructured, compound-symmetric, and autocorrelative structures. This is particularly important when analyses are based on repeated observations with a short spacing in the intervals (Schwartz & Stone, 1998). Finally, the amount of variance explained by a conditional model is estimated by comparing its residual

variances at each level with the residual variances obtained from fully unconditional models (Hox, 2002; Snijders & Bosker, 1999). The significance in variance reduction is determined by calculating the deviance reduction. The deviance (-2 Log Likelihood) is a measure of model fit that can be used to compare the relative fit of two nested models. The deviance reduction follows a chi-square distribution with degrees of freedom equal to the difference in the number of parameters of two nested models (Bryk & Raudenbush, 1992; Campbell & Kashy, 2002).