Modeling Coordination Costs Due to Time Separation in Global Software Teams

J. Alberto Espinosa and Erran Carmel
American University, Washington DC
Kogod School of Business
alberto@american.edu; carmel@american.edu

ABSTRACT

Research to date has not attempted to model coordination in global software teams. We formulate a preliminary collaboration model for a dyad to help us understand the consequences of time separation. We first describe the model and its theoretical foundations and we then evaluate the model by simulating several thousand observations and running regression models to inspect the effect of different variables on coordination costs. We then make suggestions for further extension of the model to include more complex scenarios with multiple collaborators and fewer assumptions. Our evaluation shows that the consequences of time separation are complex and that we need to understand them well before we can make claims about coordination outcomes in larger software teams that are separated by time zones.

1. INTRODUCTION

New team configurations are increasingly carried out across geographic and temporal boundaries. One recent study of a global software development organization uncovered the presence of 15 global locations, spanning 19 time zones [1]. Such complex configurations have led to interest in the effect of distance on coordination in global software teams (GSTs). Difficulties due to geographic dispersion are usually thought to correlate with time zone differences. With some exceptions [2, 3], most research has not distinguished between the two. Our objective in this paper is to present a model that represents these coordination costs in which we distinguish distance from time separation. In this article we begin to lay the groundwork for a more rigorous inquiry on this topic. We introduce a mathematical model of interdependent work between two actors and measure the coordination costs of work that moves from one actor to another. In particular, we measure these costs in all 4 conditions of the classic time-place matrix [4]. We simulate different conditions using this model.

2. TIME SEPARATION

Team members are separated by time when there are differences in working hours, time zones, and/or working rhythms that reduce the time available for same-time (i.e., synchronous) interaction. For example, teams separated on an east-west axis have fewer overlapping work hours than teams separated on a north-south axis [5], thus making it more difficult for the former to coordinate and communicate. Even co-located teams can be separated by time if their members work in different shifts.

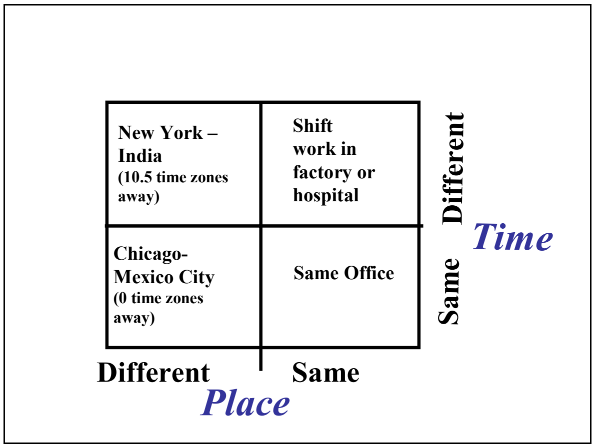


Figure 1: The time-place matrix with examples

Effective coordination is strongly influenced by effective communication since human beings communicate more effectively when in close proximity [6]. Thus, while geographic distance affects coordination, it is possible that many coordination problems are the result of time separation, which makes it difficult for members to interact synchronously. Even small time zone differences of one hour were found to bring surprising difficulties to two teams located in England and Germany [7]. And in the absence of face-to-face communications, GSTs have a menu of asynchronous and synchronous technologies to choose from. Media Richness theory [8] suggests that individuals should select communications technologies that are rich – e.g., face-to-face, video conference – but when GSTs are separated by many time zones they are forced to choose the less rich, asynchronous communication tools – e.g., e-mail, voice-mail.

The degree of task dependency plays a key role on coordination. When two team members with tightly coupled task dependencies collaborate, time zone differences can disrupt coordination. Not being able to pick up the phone and call other members can slow down a group's progress. Frequently, requests are not clear, requiring further communication. When team members are working face-to-face, the clarification may be nearly instantaneous. However, when team members are distant, clarification may introduce delay. Furthermore, unclear communication exposes the team to "vulnerability costs" – e.g., misunderstandings, rework..

On the other hand, there are cases where time zone differences may actually be beneficial. *Follow-the-sun* work [9] takes advantage of time zone differences to speed up project work. For example, a team in New York can hand off work at the end of their day to be continued by team members in India. The Indian team can progress on the task while the New York staff sleep overnight. In fact, GSTs often adjust their work to overcome time zones differences – e.g., overlap work hour windows; liaisons whose work hours are the same as the other site; batch work delivered toward the end of the day; and periodic team member travel to interact face-to-face.

3. THE MODEL

3.1 Theoretical Foundations

Coordination is "the management of dependencies" in a task. If two members can carry out a joint task independently, then there is no need to coordinate. Conversely, when two members carry out a task with tightly coupled dependencies, these dependencies need to be managed either by structuring task activities or by communicating [10-12]. We focus in this article on coordination via communication. However, coordination theory [13-15] thus far has not taken into account delays resulting from time zones differences. We try to fill this gap in this paper, beginning with a simple dyad model, influenced by Malone [13], but our model departs from his in a number of accounts:

First, Malone's model analyzes different coordination structures based on different patterns of communication and decision-making that a set of actors can use. Our model employs only two actors that need to carry out a task with tightly coupled dependencies. Second, Malone's model assumes that actors employ their production capacities optimally and that different agents have different capacities to produce. We don't make such assumption in our model because there are only two actors, one (R) who requests a task from a task provider (P) because of a dependency (i.e., R cannot continue the task until P carries out the requested task). Third, Malone's model does not incorporate time and distance separation among actors. In contrast, we specifically model time and distance separation among actors. We model distance separation as actors being either co-located or separated by distance.

3.2 Model Formulation

There are only two actors in our simple dyadic model, R and P. R has a workflow dependency with P. A single collaboration act in this context consists of the following: (1) R communicates a request to P; (2) P carries out the requested task; and (3) P communicates completion of the task to R. The model is constructed with *cost* as the dependent variable, which is, in turn, composed of three costs: (1) Production costs – due to the actual time necessary to complete the task; (2) Coordination costs – due to delay; and (3) Vulnerability costs – due to unclear messages. A message can be unclear, with some probability, which can lead to one of two conditions: (a) a request for clarification, which results in an additional cost due to delay; and (b) rework, which leads to both additional production costs and a cost of further delays.

We developed different formulas for the 8 different possible task conditions (2x2x2), depending on whether the working time overlap occurs at the beginning or end of the requestor's work day; whether the request arrives during or outside of the overlapping time; and whether the task is completed and notified during or outside of the overlapping time. For simplicity of illustration, and because the formulas don't change too much, we only analyze the model with overlapping time occurring at the end of R's work day. We model time separation based on an overlap index [5, 16] between the two actors. In Appendix A we present the resulting conditions and formulas in detail.

3.3 Assumptions

We made a number of simplifying assumptions to the model in order to test its robustness, which we can later relax to evaluate more complex collaboration models: (1) A task is composed of individual and shared portions. Actors are equally capable of doing their individual tasks. The shared portions contain dependencies that are coordinated via communication; (2) Coordination failures are due to unclear communications, creating vulnerability costs (i.e., further communication to clarify the message or re-work); (3) The probability of unclear messages increases as the richness of the communication medium used decreases. Only one clarification message is necessary to resolve unclear request messages; (4) The task is a software task and the production object is digital and it can be sent across a network in 0 time units. Similarly, messages sent synchronously or asynchronously arrive instantly; (5) There is only one synchronous and one asynchronous link between R and P; (6) The task is high priority and time constrained; (7) Non face-to-face communication is conducted electronically, and when working hours overlap actors prefer to communicate synchronously (e.g., telephone, video conference); they communicate asynchronously (e.g., e-mail, shared databases) otherwise; (8) All tasks requested by R are immediately accepted and carried out competently by P and there is no parallel multi-tasking. Once P has full information about the requested task, P's production costs are the same regardless of time or distance separation; (9) Time is measured from R's perspective. If P is processing a task during R's non-work hours it has no time delay consequences for R.

	Coordination Costs				Vulnerability Costs			
	Main Effects		+ Interaction		Main Effects		+ Interaction	
Variable	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
Constant	**-390.64**	**<0.001**	**-409.64**	**<0.001**	**-55.18**	**<0.001**	**-69.65**	**<0.001**
Request Time	**-353.81**	**<0.001**	**-49.51**	**<0.001**	**-133.62**	**<0.001**	**-27.13**	**<0.001**
Task Duration	**721.50**	**<0.001**	**942.53**	**<0.001**	**71.88**	**<0.001**	9.80	0.143
Overlap Index	**-159.64**	**<0.001**	-2.51	0.492	**-135.13**	**<0.001**	0.21	0.923
Distributed	**594.14**	**<0.001**	**600.09**	**<0.001**	**24.71**	**<0.001**	**26.82**	**<0.001**
Time Separated	**208.39**	**<0.001**	**205.34**	**<0.001**	**70.91**	**<0.001**	**75.21**	**<0.001**
Distributed & Time Separated	**749.31**	**<0.001**	**759.58**	**<0.001**	**123.90**	**<0.001**	**122.93**	**<0.001**
ReqTime x TskDur			**-321.26**	**<0.001**			**-26.33**	**0.020**
ReqTime x Overlap			**471.20**	**<0.001**			**255.99**	**<0.001**
TskDur x Overlap			**496.89**	**<0.001**			**-28.37**	**0.009**
ReqTime x Distr			5.87	0.279			4.79	0.136
ReqTime x TimeSep			**-610.74**	**<0.001**			**-170.78**	**<0.001**
ReqTime x Distr&Time			**-608.48**	**<0.001**			**-257.61**	**<0.001**
TskDur x Distr			1.52	0.922			3.87	0.676
TskDur x TimeSep			**-416.15**	**<0.001**			**87.36**	**<0.001**
TskDur x Distr&Time			**-454.54**	**<0.001**			**154.38**	**<0.001**
Overlap x Distr			*9.32*	*0.071*			4.14	0.177
Overlap x TimeSep			**-366.62**	**<0.001**			**-235.66**	**<0.001**
Overlap x Distr&Time			**-262.92**	**<0.001**			**-305.90**	**<0.001**
R-sq	0.854		0.974		0.571		0.901	
R-sq Change			0.120				0.330	
R-sq Change P-Value			<0.001				<0.001	

Table 1: Regression Analysis Results

4. MODEL EVALUATION

We evaluated the robustness of the model with a simulation of 11,000 observations and then exploring the effect of the timing of requests, task duration, and time overlap on coordination and vulnerability costs. Again, we only discuss the case when this overlap occurs at the end of the requestor's work day. The request time (Rt) variable was generated randomly from a uniform distribution (0,1), with 1 being a full work day. The task duration (Tt) variable was also generated randomly from a normal distribution with an average of 0.25 (1/4 of a work day) and a standard deviation of 0.1. We fixed all other parameters as follows (see Appendix A): Cla=\$100 and Cls=\$500 per day; Cma=\$10 and Cms=\$50 per message; Cd=\$1,000 and Cp=\$1,000 per day. While these costs are arbitrary, they serve the purpose of helping illustrate and evaluate the model. Further evaluations of this model will incorporate variable costs along a normal distribution.

The probability that a request was unclear was fixed at 10%, 30%, 50% and 70% for the four conditions, respectively, face-to-face, distributed, time-separated and separated by time and distance. These probabilities are arbitrary but based on our expectation that, as the richness of the communication media diminishes the probability of unclear messages increases. The probability differences are purposefully wide at this point to amplify their impact on coordination costs and make these effects more noticeable. Also, there is a probability of 30% that unclear messages will lead to re-work and 70% that it will simply lead to a request for further clarification with no re-work. If re-work is necessary, it is assumed that, on average, 30% of the work completed will have to be redone, thus increasing production costs. Finally, we assume one request per day.

Regression results from Ordinary Least Squares models suggest that the model is robust and that it behaves as expected. Both regression models were run first including only main effects and then adding interaction variables. The results are presented in Table 1. The inclusion of interaction variables had a significant increase in explained variance (R^2) in both models, suggesting that we need to pay attention to these interactions. The regression models with main effects (only) yielded intuitive and similar results for coordination costs and vulnerability costs. Both costs increase with longer tasks, and with time and/or distance separation. Both costs decrease when requests come later in the day (i.e., closer to overlapping hours) and when there are more overlapping hours in the day. The coefficients for coordination costs are larger in absolute value for coordination costs than vulnerability costs, but this difference will change as we change cost parameters in the future.

Most main effects remained significant and retained their signs when the interaction variables were added to both models with a few exceptions. The main effect of overlapping hours became non-significant in both models and the main effect of task duration became non significant in the vulnerability costs model. The sign and significance levels of the interaction coefficients for the overlap variable indicate that the amount of overlapping time has a significant effect for teams that are separated by time or by time-and-distance, as one would expect, but, naturally, it does not have an effect on face-to-face and distributed-same time conditions. This is an expected result since these teams have full overlap in their working hours.

The negative interaction coefficients for task request time and time separation suggest that issuing task requests closer to overlapping working hours reduces coordination costs when actors are separated by time, which is also an intuitive result. Interestingly, task duration increases coordination costs, as expected, but this effect diminishes with time separation, because the task provider T can work during the requestor R's off work hours. This is consistent with the benefits of *"follow-the-sun"* noted in Section 2.

The negative interaction coefficient of request time with task duration suggest that issuing task requests closer to overlap time reduces coordination and vulnerability costs, but more so for tasks of longer duration. On the other hand, the positive interaction coefficient of request time with work overlap time suggests that the beneficial effects of issuing task requests later in the day, closer to overlap time are, naturally, offset as the overlap working hours get larger. In time-separated work contexts, request timing is critical in reducing coordination and vulnerability costs, but this becomes less important with less time separation.

Finally, the interaction coefficient of task duration with overlap time was positive for coordination costs and negative for vulnerability costs. This suggests that coordination costs increase with task duration, especially when there is less time separation (i.e., more overlap), but this is somewhat offset by lower vulnerability costs because it is less costly to clarify miscommunication when there is less time separation.

5. DISCUSSION AND FUTURE RESEARCH

As expected, coordination and vulnerability costs are very sensitive to the time at which the task is requested in time separated contexts, but this effect diminishes with less time separation with, of course, no effect with full overlap (i.e., co-located or distance separated only). This and the other results discussed in the prior section suggest that the model we have formulated in this paper is well behaved. Collaborations in which more than two actors are separated by time are much more complex than the model we have presented here. But the robustness of our dyadic model gives us confidence that this model can be extended to more complex coordination structures. We plan to expand the model by varying cost and operational parameters to simulate different realities. For example, delay costs are higher where time-to-market is critical. On the other hand, production costs may be much higher in situations where software production requires expensive resources (e.g., sophisticated testing labs, scarce expertise). Other interesting manipulations include: giving actors a choice of synchronous and asynchronous communication facilities, each with different costs, and then evaluate the tradeoffs of reducing coordination and vulnerability costs, against communication costs; assessing new collaboration tools by reducing the probability of unclear messages due to better communication effectiveness; actors could also reflect different production costs (as is typical today with offshore work). In sum, we plan to expand our model by progressively relaxing assumptions.

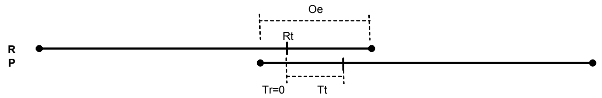
6. REFERENCES

1. Orlikowski, W., *Knowing in Practice: Enacting a Collective Capability in Distributed Organizing.* Organization Science, 2002. 13(3): p. 249-273.
2. Espinosa, J.A., et al., *Team Boundary Issues Across Multiple Global Firms.* Journal of Management Information Systems, Spring 2003. 19(4)
3. Watson-Manheim, M.B., K. Crowston, and K.M. Chudoba. *A New Perspective on Virtual: Analyzing Discontinuities in the Work Environment, 35th. Hawaiian International Conference on System Sciences*. HICSS 2002. Big Island, Hawaii: IEEE.
4. Bullen, C. and J. Bennett, *Groupware in Practice: An Interpretation of Work Experiences*, in *Groupware and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration*, R. Baecker, Editor. 1993, Morgan Kaufman Publishers: CA. p. 69-84.
5. O'Leary, M.B. and J.N. Cummings. *The Spatial, Temporal, and Configurational Characteristics of Geographic Dispersion in Teams*. *Academy of Management Conference,* 2002. Denver, Co.
6. Kiesler, S. and J.N. Cummings, *What Do We Know About Proximity in Work Groups? A Legacy of Research on Physical Distance*, in *Distributed Work*, P. Hinds and S. Kiesler, Editors. 2002, MIT Press: Cambridge, MA. p. 57-80.
7. Grinter, R.E., J.D. Herbsleb, and D.E. Perry. *The Geography of Coordination: Dealing with Distance in R&D Work, ACM SIGGROUP Conference Work (Group 99)*. 1999. Phoenix, AZ: ACM Press.
8. Daft, R. and R. Lengel, *Organizational Information Requirements, Media Richness and Structural Design.* Management Science, 1986. 32(5).
9. Carmel, E., *Global Software Teams*. 1999, Upper Saddle River, NJ: Prentice Hall.
10. March, J. and H. Simon, *Organizations*. 1958: John Wiley and Sons.
11. Thompson, J., *Organizations in Action*. 1967: McGraw-Hill.
12. VanDeVen, A.H., L.A. Delbecq, and R.J. Koenig, *Determinants of Coordination Modes Within Organizations.* American Sociological Review, 1976. 41(April): p. 322-338.
13. Malone, T., *Modeling Coordination in Organizations and Markets.* Management Science, 1987. 33(10): p. 1317-1332.
14. Malone, T. and K. Crowston. *What is Coordination Theory and How Can it Help Design Cooperative Work Systems*. in *Computer Supported Collaborative Work.* 1990: ACM Press.
15. Malone, T. and K. Crowston, *The Interdisciplinary Study of Coordination.* ACM Computing Surveys, 1994. 26(1): p. 87-119.
16. O'Leary, M.B. *Varieties of Virtuality: Separate but not Equally, FIU Workshop on Distributed Work and Virtuality*. 2001. Miami, Fla.

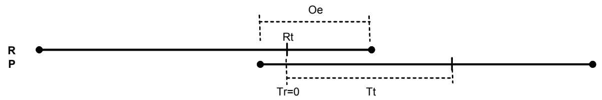
Appendix A: Model Variables and Formulas

When overlapping working hours are at the end of R's workday

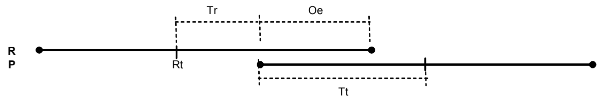
1. R(Oe)C(Oe) Request comes during overlapping hours and is also finished during overlapping hours



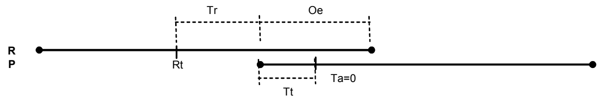
2. R(Oe)C(after-Oe) Request comes during overlapping hours, but is finished after overlapping hours



3. R(bef-Oe)C(after-Oe) Request comes before overlapping hours and is finished after overlapping hours



4. R(bef-Oe)C(Oe) Request comes before overlapping hours and is finished during overlapping hours



Variables:

λ = number of tasks per day that R requests from P
Cp = P's production cost (per time unit)
Cd = R's cost of delay (waiting) for task to be processed and getting confirmation
Cl = Cls + Cla = Cost per day of maintaining a (synch + asynch) connection between R and P
Cm = Cms + Cma = Cost of sending a (synch + asynch) message from R to P or viceversa
Ob = Overlap index (0-1), proportion of overlap time in work hours, at beginning of R's work day
Oe = Overlap index (0-1), proportion of overlap time in work hours, at end of R's work day
Rt = Time when R requested task from P, expresed as a fraction (0-1) of time from the start of the work day
Tt = Task duration time or time it takes P to complete the requested task, as a fraction (0-1) of the work day
Td = Delay time between R's request and P's acknowledgement, measured from R's perspective as a fraction
Pn = Probability that a message is not clear; varies with collaboration mode: Pn(TP)>Pn(T)>Pn(P)>Pn(O)
where, TP=separated by time and place, T=by time only, P=by place only, O=face-to-face

O. SameTime-SamePlace -- Co-located

Assumptions: Cl = 0; Cm = 0; Tr=0; Ta=0; Pn(O) is very low
Production Costs = λCpTt (same across all 4 collaboration modes O, P, T and TP
Coordination Costs = λTtCd
Vulnerability Costs = λPn(O)[(Pr)(RwCpTt)+(1-Pr)(0+0)(Cd)] = λPn(O)[(Pr)(RwCpTt)]; Pn(O) is very low, on average

P. Same Time-Different Place -- Distributed Synchronous (separated by place)

Assumptions:
Cla = 0 or Cl = Cls; Cma = 0 or Cm = Cms; T2=0; T5=0
Pn(TP)>Pn(P)>Pn(O); probability of unclear messages is higher when separated by more boundaries
Production Costs = λCpTt
Coordination Costs = Cls + 2λCms + λTtCd
Vulnerability Costs = λPn(P)[(Pr)(RwCpTt)+(Pf)(2Cms+0+0)] = λPn(P)[(Pr)(RwCpTt)+(1-Pr)(2Cms)]

T and TP. Different Time-Same Place/Different Place -- Asynchronous (Co-Located and Distributed)
With Overlap at the End of R's workday (Oe)

1. R(Oe)C(Oe) Request comes during overlapping hours and is also finished during overlapping hours

(T) Td = Tt; Tr=0; Ta=0; Cm=0 -- (PT) same, except that ***Cm=Cms***
Coordination Costs (T) = Cla + λTtCd -- (PT) = Cla + ***Cls*** + ***2λCms*** + λTtCd
Vulnerability Costs (T) = λOePn(O)(Pr)(RwCpTt) -- (PT) = λ***Pn(P)***(Pr)(RwCpTt)

2. R(Oe)C(after-Oe) Request comes during overlapping hours, but is finished after overlapping hours

Td = 1-Rt; Tr=0; Ta=1-Rt
(T) Cm=Cma (for both request and acknowledgement) -- (PT) Cm = Cma + ***Cms***
Coordination Costs (T) = Cla +λCma + λ(1-Rt)Cd
(PT) = Cla + ***Cls*** + λ(Cma+***Cms***) + λ(1-Rt)Cd
Vulnerability Costs (T) = λPn(O)[(Pr)RwCpTt+(1-Pr)(Cma+(1-Rt)Cd)
(PT) = λ***Pn(P)***[(Pr)RwCpTt+(1-Pr)(Cma+***Cms***+(1-Rt)Cd)

3. R(bef-Oe)C(after-Oe) Request comes before overlapping hours and is finished after overlapping hours

Td = 1-Rt; Tr = 1-Rt-Oe; Ta=0
(T) Cm=Cma (for both request and acknowledgement) -- (PT) Cm = Cma + ***Cms***
Coordination Costs (T) = Cla +2λCma + λ(1-Rt)Cd
(PT) = Cla + ***Cls*** + 2λCma + λ(1-Rt)Cd
Vulnerability Costs (T) = λPn(T)[(Pr)RwCpTt+(1-Pr)(2Cma+(1-Rt)Cd)
(PT) = λ***Pn(PT)***[(Pr)RwCpTt+(1-Pr)(2Cma+(1-Rt)Cd)

4. R(bef-Oe)C(Oe) Request comes before overlapping hours and is finished during overlapping hours

Td = 1-Rt-Oe+Tt; Tr = 1-Rt-Oe; Ta=0
(T) Cm=Cma (for request only) -- (PT) Cm=Cma+***Cms***
Coordination Costs (T) = Cla +λCma + λ(1-Rt-Oe+Tt)Cd
(PT) = Cla + ***Cls*** + λ(Cma+***Cms***) + λ(1-Rt-Oe+Tt)Cd
Vulnerability Costs (T) = λPn(T)[(Pr)RwCpTt+(1-Pr)(Cma+(1-Rt-Oe)Cd)]
(PT) = λ***Pn(PT)***[(Pr)RwCpTt+(1-Pr)(Cma+***Cms***+(1-Rt-Oe)Cd)]