

MEASURING THE DEGREE OF INTERNATIONALIZATION OF A FIRM

Daniel Sullivan*
University of Delaware

Abstract. In spite of both positivistic and instrumental research, the reliability of measuring the degree of internationalization of a firm remains speculative. We collected data on nine attributes of seventy-four American manufacturing MNCs. Alpha, factor, and frequency analyses revealed a linear combination of five variables with a reliability coefficient of .79 as a measure of the degree of internationalization of a firm. We discuss the statistical and conceptual properties of the scale and their implications for content and construct validity.

The validation of theories of international business has not matched the robustness of their development. Confirmation has been hindered by the lack of reliable measures, the ensuing inability to disentangle the distorting influences of measurement error, and ultimately, the impossibility of establishing content and construct validity. The absence of a coherent approach to establish the validity of measurements results in empirical investigations that are disjointed and inconclusive, a proliferation of partially tested or untested propositions, and a segregation of the theory-building process from the hypothesis-testing phase of research. Consequently, we are unable to create a cumulative structure of theoretical, derived and empirical concepts that provide purpose to subsequent studies.

For example, despite its theoretical and practical centrality, estimating the degree of internationalization (DOI) of a firm remains arbitrary. Some scholars, relying on the loosely structured or even unstructured inductive frameworks of the instrumental approach, try to infer the DOI of a firm by examining the evolution, structure, and processes of relationships among its demographic, strategic, market, organizational, product, and attitudinal characteristics of international expansion [Johanson and Vahlne 1977; Forsgren 1989; Welch and Luostarinen 1988]. The moot external validity of gestalt-type measures has proved troublesome. Such measures do not document the reliability of measurement and, in turn, the validity of interpretation [Phillips and Bagozzi 1985]. This method tends to inflate or deflate the observed

*Daniel Sullivan studies strategic and organizational features of the MNC as well as principles of internationalization theory. In addition, he has been awarded numerous teaching honors from undergraduate, MBA, and Executive MBA students.

The author thanks Professor Alan Bauerschmidt of the University of South Carolina for initial inspiration.

Received: February 1992; Revised: August 1992, April & October 1993; Accepted: November 1993.

association among variables, thereby increasing the odds of accepting findings despite their contradiction of received theory [Nunnally 1978]. Instrumental research is prone to more than the usual amount of random error given that informants, researchers, or assistants must make inferences about macro phenomena, presume actors' motivations, and perform aggregations of tasks and events [Seidler 1974]. When a single researcher makes the call, analyses are susceptible to systematic error due to his or her interpretation of the process [Ericsson and Simon 1980]. Finally, in both cases, reconstructing the often idiographic sequence of internationalization requires pinpointing many events and relationships; failure to specify, represent, and control for random and systematic error in measurement can lead to inconsistent estimates [Cook and Campbell 1979]. The consequence of these deficiencies is exemplified by the contradictory reports regarding the sequence of internationalization, as postulated by Johanson and Vahlne [1977]. Cavusgil and Godiwalla [1982:53] conclude that "studies of British, Swedish, French, and American firms all support a gradual pattern in internationalization." Others report that the thesis of incremental internationalization fails to explain the phenomenon [Millington and Bayliss 1990; Sullivan and Bauerschmidt 1990; Turnbull 1987].

In search of greater precision, some scholars apply positivistic principles and try to differentiate DOI by blocking on a single criterion. These researchers study the relationship between single independent and dependent variables, relying on deductive frameworks to structure analysis. Proxies of DOI include foreign subsidiaries' sales as a percentage of total sales [Stopford and Dunning 1983], foreign assets as a percentage of total assets [Daniels and Bracker 1989], and number of foreign subsidiaries [Stopford and Wells 1972]. Granted, a single item facilitates replication. However, the medley of measures has neither helped establish a standard criterion nor clarified the content validity of measurement. These outcomes are inevitable. A single item does not permit one to take measurement error into account in analyses [Campbell and Fiske 1959; Schoenfeldt 1984]. In terms of testing hypotheses, the impossibility of determining the reliability of a single measure increases the probability of a Type I or Type II error [Bagozzi, Youjae and Phillips 1991]. Nunnally [1978] notes that using a single measure for a concept creates the risk that the measure will be confounded by existing method biases. Also, employing a common method to derive measures of independent and dependent variables can artificially increase the association observed therein [Bollen 1989; McDonald and Marsh 1990]. Furthermore, because a single item represents only a limited portion of the domain, it tends to misrepresent the construct. Similarly, when only a single item—or even just a single aspect of a multi-attribute domain—is used, any unusual circumstances that might distort the validity of the measure will contaminate, if not ruin, the results [Nunnally 1978]. Finally, studies utilizing a single item often use parametric statistical techniques but neglect many of the controls for method variance that Mitchell [1985] emphasizes to ensure consistency with their presumption of an absence of measurement error.

Consequently, we often lack the statistical evidence to support the premise that measurements and concepts correspond. Buckley and Casson [1983:201] as well as Geringer, Beamish and daCosta [1989:118] illustrate this limitation; reporting results gained by using a single-item measure of DOI, the former attribute the “mixed results” of the effects of the internationalization of a firm to the problems of accurate measurement of DOI, while the latter caution that their single-item indicator of “the degree of internationalization represents a rough measure, and future research may improve upon its operationalization.” Thus, the seemingly routine caveat regarding possible, if not probable, measurement error in positivistic research underscores the risk that a single-item measure may result in spurious confirmation or distorted estimates of the relationships allegedly constituting internationalization.

A Case in Point

Some may counter that the preceding contradictions and conditionalities are the by-product of research that has not yet reached critical mass. A detailed illustration rejects this supposition. An elemental issue of international business is whether diversifying internationally improves the financial performance of a firm. A priori, the practices of thousands of companies indicate yes. However, looking to the literature for confirmation proves futile. We categorized seventeen empirical studies of the relationship between the DOI of a firm and its financial performance on the basis of whether the study found a positive, indeterminate, or negative relationship. Table 1 shows that six studies reported a positive, six an indeterminate, and five a negative relationship. The theoretical clarity of the relationship between DOI and financial performance makes such empirical disarray disturbing. Searching for an explanation prompted evaluating the design of these seventeen studies. Table 2 shows that great care had been taken to study representative and diverse firms and develop meaningful measures of financial performance. However, glaring in its consistency is the inevitable use of Foreign Sales as a Percentage of

TABLE 1
The Reported Direction of the Relationship between Degree of Internationalization and the Financial Performance of the Firm

Positive	Indeterminate	Negative
Vernon [1971]	Horst [1973]	Siddharthan & Lall [1982]
Dunning [1985]	Hughes, Logue & Sweeny	Kumar [1984]
Grant [1987]	[1975]	Michel & Shaked [1986]
Grant, Jammine &	Buckley, Dunning & Pearce	Shaked [1986]
Thomas [1988]	[1977]	Collins [1990]
Daniels & Bracker [1989]	Rugman, Lecraw & Booth [1985]	
Geringer, Beamish &	Yoshihara [1985]	
daCosta [1989]	Buhner [1987]	

TABLE 2
Summary of Empirical Studies of the Relationship between
Financial Performance and the Degree of Internationalization

Study	Sample	Measure of DOI	Measure of Performance	Report
Vernon [1971]	The Fortune 500 in 1964	<i>FSTS</i>	ROS ROA	MNCs earned higher ROS and higher post-tax ROA than did non-MNCs.
Horst [1973]	1191 United States Industrial firms	<i>FSTS</i>	Net profits	After controlling for firm size, unable to discriminate MNCs from non-MNCs on basis of net profits.
Hughes, Logue & Sweeny [1975]	46 U.S. MNCs 50 U.S. non-MNCs	<i>FSTS</i>	<i>Beta</i> ; Risk-adjusted returns	MNCs had higher risk-adjusted returns to shareholders, but lower average returns to shareholders as well as a lower average <i>beta</i> .
Siddharthan & Lall [1982]	The 500 and 100 largest U.S. and non-U.S. MNCs, respectively, in 1972	<i>FSTS</i>	Firm growth	Controlling for firm size, advertising intensity, R&D intensity, scale economies, and profitability indicates that DOI had a negative effect on the rate of firm growth.
Kumar [1984]	672 British firms, 1972-1976	<i>FSTS</i>	ROS, ROA	MNCs outperformed non-MNCs for both ROS and ROA; however, OLS revealed negative but nonsignificant relationship between <i>FSTS</i> and profits.
Buckley, Dunning & Pearce [1977]	The 636 and 866 largest MNCs of the world, respectively, for 1972 and 1977	<i>FSTS</i>	ROA	Results are "somewhat inconsistent"; while statistically significant in 1972 for full sample, relationship was insignificant for the full and U.S. sample in 1977.
Dunning [1985]	188 large British MNCs, 1979	<i>FSTS</i>	ROS	Degree of multinationality had a positive but statistically insignificant relationship with ROS.
Yoshihara [1985]	The largest 118 Japanese firms	<i>FSTS</i>	ROE	MNCs outperformed non-MNCs in terms of ROS but the differences were not statistically significant.
Rugman, Lecraw & Booth [1985]	The 50, 50, 20, 10, and 24, respectively, largest U.S., European, Japanese, Canadian, and Third World MNCs	<i>FSTS</i>	ROE	No evidence of excessive profits for MNCs. Most MNCs had ROE between 10%-14%, which were similar to the ROE of domestic firms of similar size.

TABLE 2
(continued)

Study	Sample	Measure of DOI	Measure of Performance	Report
Shaked [1986]	58 U.S. MNCs, 43 non-U.S. MNCs for 1980-1982	<i>FSTS</i>	ROA, <i>Beta</i>	ROA did not differ significantly between MNCs and domestic firms. MNCs had lower systematic risk, odds of insolvency, and equity variability, but also experienced lower sales growth than domestic firms.
Michel & Shaked [1986]	58 U.S. MNCs and 43 non-U.S. MNCs for 1973-1982	<i>FSTS</i>	Risk-adjusted returns	Risk-adjusted returns to stockholders were higher for domestic firms than for MNCs.
Grant [1987]	304 British firms for 1968-1984	<i>FSTS</i>	Sales growth; ROS, ROA, ROE	Internationalization was positively associated with superior profitability over a period of thirteen years.
Buhner [1987]	40 West German firms	<i>FSTS</i>	Risk-adjusted returns, ROE, ROA	Single business firms with high degrees of internationalization outperformed other firms. Performance between others was not significantly different.
Grant, Jammine & Thomas [1988]	304 British firms for 1972-1984	<i>FSTS</i>	ROA	No evidence of an association between high levels of DOI and a significant downturn in profitability.
Daniels & Bracker [1989]	116 U.S. MNCs	<i>FSTS</i> ; <i>FATA</i>	ROS, ROA	Performance improved significantly as <i>FSTS</i> and <i>FATA</i> increased to 50%. While not a significant association, performance seemed to fall as <i>FSTS</i> and <i>FATA</i> increased beyond this threshold.
Geringer, Beamish & daCosta [1989]	The largest 100 U.S. and 100 European MNCs of 1981	<i>FSTS</i>	ROS, ROA	Financial performance improved monotonically as <i>FSTS</i> grew but peaked and declined when <i>FSTS</i> crossed the revealed threshold of 60%-80% of internationalization.
Collins [1990]	150 firms of the Fortune 500	<i>FSTS</i>	Total risk, Leverage, <i>Beta</i>	Performance of MNCs with developed country DFI was equal to domestic firms but the performance of MNCs with operations in developing countries was inferior to other MNCs and domestic firms.

Total Sales (FSTS) as the sole estimator of DOI; each of the seventeen studies designated this ratio as the independent variable. By no means did we selectively include only studies that set *FSTS* as the independent variable. Selection was more random than not. Indeed, including less visible research would have only further dramatized that the vast majority of empirical studies of DOI and financial performance estimated the former construct with a measure that is intrinsically unreliable and has, at best, speculative validity. The consequence of the consistency of using *FSTS* is stark: we are unable to state with certainty that international diversification will improve the financial performance of a firm.

THE RESEARCH ISSUE

Fiske [1982:82] charged that method variance in social science research is “pervasive, ubiquitous.” Phillips and Bagozzi [1985] and Bagozzi, Youjae and Phillips [1991:405] concurred, reporting that among extant studies in organizational research, method variance is “not only prevalent, but relatively large.” More precisely, Schmidt and Hunter [1981] reported that up to 72% of the variance in correlational studies is due to measurement and sampling error. There is little cause, given the contradictions and caveats noted above as well as the implications of Tables 1 and 2, to believe that international business research is immune to this virus. This paper tries to address this flaw. It describes an effort to improve the content validity of measuring the DOI of a firm. Two parameters frame our efforts: one, *a priori*, the eventual measure tries to integrate the instrumental and positivistic methods and, two, the eventual measure only uses archival data in order to facilitate replication.

RESEARCH METHODS

Sample

Since 1979, *Forbes* has annually ranked the “Most International” 100 American manufacturing and service firms on the basis of total foreign revenues. From 1979 to 1990, the *Forbes* rankings included 119 manufacturing companies. The sample for this study consisted of 74 manufacturing firms from this set.¹

Undoubtedly, this group of MNCs had some sampling and criterion bias. We believe the former was slight given that these 74 MNCs are representative of the type of company that will be studied with the eventual measure. Chandler [1986:409] offers corroboration. His analysis of the evolution of the MNC found that companies in seven industries—food, chemical and allied products, pharmaceuticals, petroleum and refining, industrial machines, electronics, and transportation—have accounted for the majority of U.S. MNCs since the 1880s.² More precisely, he reported that approximately 74% of the 200 largest U.S. MNCs consistently were from these industries. These seven industries contribute 83% of our sample.

Concerning criterion bias, *Forbes* ranks MNCs according to the absolute level of foreign revenues. Ranking MNCs according to another criterion

would change the composition of the sample. However, as did Buckley, Dunning and Pearce [1977], Beamish and daCosta [1984], Stopford and Dunning [1983], and Daniels and Bracker [1989], we accept the premise that a company's foreign sales are a meaningful first-order indicator of its involvement in international business.

Research Variables

Nunnally [1978:258] concluded that "since content validity depends on a rational appeal to an adequate coverage of important content, an explicit outline of content provides a basis for discussing content validity." The literature suggests that the DOI of a firm has three attributes: *performance* (what goes on overseas, Vernon [1971]), *structural* (what resources are overseas, Stopford and Wells [1972]), and *attitudinal* (what is top management's international orientation, Perlmutter [1969]). Integrative works emphasize that DOI is not an absolute state but a continuous choice that managers make relative to domestic circumstances [Forsgren 1989; Welch and Luostarinen 1988]. Therefore, all measures are ratios. As noted earlier, there is no shortage of potential measures. However, our rule of only using archival data restricted including and, if need be, designing measures that we could estimate objectively.

Five measures operationalize the performance attribute of internationalization. A common measure of DOI is Foreign Sales as a Percentage of Total Sales (*FSTS*) [Daniels and Bracker 1989; Geringer, Beamish and daCosta 1989; Stopford and Dunning 1983]. Caves [1982:198] notes that "R & D activities themselves predict the rise of MNEs," a report extended by Franko's [1989:470] finding that Research and Development Intensity (*RDI*) was a "principal, perhaps *the* principal, means of gaining market share in global competition"; thus we included this variable. Similarly, Caves [1982:10], Capon, Farley and Hoening [1990], and Keown, Synodinos, Jacobs and Worthley [1989] suggest that the scale of an MNC's marketing function, routinely operationalized by Advertising Intensity (*AI*), helps explain a firm's international involvement. Sullivan and Bauerschmidt [1989] report that the degree of export activity—Export Sales as a Percentage of Total Sales (*ESTS*)—discriminates the relative internationalization of American and European firms. Finally, Eppink and Van Rhijin [1988] suggest estimating DOI with Foreign Profits as a Percentage of Total Profit (*FPTP*).

Two measures operationalize the structural attribute of DOI. Daniels and Bracker [1989] used Foreign Assets as a Percentage of Total Assets (*FATA*) to estimate the material international character of a firm. Similarly, Stopford and Wells [1972] and Vernon [1971] reason that the number of foreign subsidiaries distinguishes the international involvement of a firm. Variability in the scale and scope of subsidiaries among MNCs prompted standardizing this index by gauging the number of foreign subsidiaries relative to the company's total number of distinct operating units for a measure of Overseas Subsidiaries as a Percentage of Total Subsidiaries (*OSTS*).

Measuring the attitudinal attribute of DOI is difficult in that archival data do not inventory the psychometric attributes of managers or firms. While the empirical literature offers no standards, conceptual works propose proxies. Perlmutter [1969] and Maisonrouge [1983] reason that a top manager's international orientation correlates positively with the extent of his or her international experience. Thus, we began measuring Top Managers' International Experience (*TMIE*) by tallying the cumulative duration of top managers' international assignments, as identified by the firm and as summarized in each manager's company-reported career history. We then weighted this sum by the reported total number of years of work experience of the top management team of the firm as identified by the firm.³

The second attitudinal measure tries to estimate the psychic dispersion of the international operations of a firm. Elaborating the notion of bounded rationality [March and Simon 1958], Johanson and Vahlne [1977] report that the scope of interaction between home and host nationals, in terms of the degree of psychic distance, correlates positively with the internationalization of the firm.⁴ Sullivan and Bauerschmidt [1990] operationalize psychic distance by estimating the geographical dispersion of the overseas subsidiaries of a firm. We follow the same logic, but estimate the Psychic Dispersion of International Operations (*PDIO*) by calibrating the dispersion of the subsidiaries of a firm among the ten psychic zones of the world as identified by Ronen and Shenkar [1985].⁵ Each zone has, as Hofstede [1993:84] and Adler, Doktor and Redding [1986] suggest, a unique "cognitive map" of the principles of management. Therefore, we presumed that the greater the dispersion of an MNC's subsidiaries across these ten zones, the greater the psychic dispersion of its international operations.

Data Sources

We calculated *FSTS*, *RDI*, *AI*, *ESTS*, *FPTP*, and *FATA* with data obtained from PC-Compustat and Annual Reports. To estimate *TMIE*, we used data from the career histories of top managers as reported in Dun's *Reference Book of Corporate Managements* [1991]. Dun's *Directory of American Corporate Families and International Affiliates, Volumes I and II* [1991] supplied the frequency and distribution of subsidiaries needed to calibrate *OSTS* and *PDIO*. We calculated a three-year average, from 1988 to 1990, for *FSTS*, *FATA*, *RDI*, *AI*, *ESTS*, and *FPTP* to control for changes in exchange rates and accounting standards. The comparatively slower rate of change for *OSTS*, *TMIE* and *PDIO* along with data availability limitations led to a single value per attribute for each MNC as of 1990.

Data Analysis

Nunnally [1978:257] notes that "although there are competing models for special problems of measurement, there is no general competitor to the linear model." The superiority of the linear model results from the capacity

of a linear combination of single items to reduce measurement error, control for confounds, and enable estimating reliability [Bagozzi, Youjae and Phillips 1991]. Therefore, we used Nunnally's [1978] "item-total analysis for constructing homogenous measures" method to estimate the internal consistency of a linear combination of single measures of DOI.⁶ This technique correlates each potential item with the item-corrected scale score, the latter defined as the sum of the scores of all other items. "Good" items were ones that correlated significantly with the scale score. In this study, a "good" item was one whose correlation with the scale exceeded .291, the .01 significance threshold from the critical points for testing zero correlation in a bivariate normal distribution. Ideally, this method identifies a set of measures that as a scale have a high *alpha* coefficient and as a cluster load on a single factor. If so, the final check is the normality of the distribution of the values of the sampled MNCs on the revealed measurement scale. To enrich analysis, we also evaluated two pertinent concerns: the comparative robustness of a multi-item scale relative to that of the conventional single-item measure as well as the degree of criterion bias.

RESULTS

Among the seventy-four companies comprising the sample, no one type of firm dominated: there were three single product, eighteen dominant product, twenty-five related technology product, ten related marketing product, seven related marketing and technology, and eleven unrelated product companies.⁷ On average, each firm participated in 19 different four-digit SIC segments, with a low of 5 and a high of 51. The sample was biased toward large MNCs: for 1990, the average *Fortune 500* rank was 53; 63 firms ranked among the *Fortune 100*; and ranks ranged from 1 to 152.

We applied the item-total analysis for constructing homogenous measures method to the nine variables. The resulting scale's reliability, as represented by the *alpha* coefficient, was .58. Two conditions pressed for additional analysis: first, the item-total correlation of *ESTS* ($r_{is} = -.09$), *AI* ($r_{is} = .07$), *RDI* ($r_{is} = .04$), and *FPTP* ($r_{is} = .02$) fell below the threshold of $r_{is} \geq .291$ and, second, Nunnally [1978:245] advises that "in the early stages of research, reliabilities of .70 or higher will suffice." Further testing found that the scale composed of *FSTS* ($r_{is} = .30$), *FATA* ($r_{is} = .41$), *OSTS* ($r_{is} = .32$), *TMIE* ($r_{is} = .37$), and *PDIO* ($r_{is} = .42$) attained the highest reliability with an *alpha* of .79. We labelled this combination the "Degree of Internationalization Scale" (*DOI_{INTS}*) (see Table 3).

Factor analysis showed that a single factor comprised *DOI_{INTS}* (see Table 4). The scree and eigenvalue of the factor corroborated the structure. The factor loadings precluded any ambiguity about the relationship between the factor and measures. The communality for each of the items, except perhaps that of *TMIE*, indicated that each explained a high proportion of the variance in the variable.

TABLE 3
Descriptive Statistics and Correlations for the Research Variables

	Mean	SD	<i>FSTS</i>	<i>FPTP</i>	<i>RDI</i>	<i>AI</i>	<i>ESTS</i>	<i>FATA</i>	<i>OSTS</i>	<i>TMIE</i>	<i>PDIO</i>
<i>FSTS</i>	.361	.14	1.00								
<i>FPTP</i>	.472	.84	.91	1.00							
<i>RDI</i>	.041	.17	-.01	-.11	1.00						
<i>AI</i>	.036	.21	-.01	.02	-.03	1.00					
<i>ESTS</i>	.073	.56	-.14	-.18	.30	-.20	1.00				
<i>FATA</i>	.318	.14	.87	.80	-.01	.07	-.20	1.00			
<i>OSTS</i>	.552	.19	.61	-.18	.53	.29	.22	.48	1.00		
<i>TMIE</i>	.095	.08	.32	.11	-.01	.21	-.01	.27	.14	1.00	
<i>PDIO</i>	.742	.18	.46	-.003	.21	.06	-.06	.34	.45	.35	1.00

TABLE 4
Final Statistics for Factor Analysis of Degree of Internationalization Scale

Factor	Variable	Loading	Communality
1	<i>OSTS</i>	.9228	.8515
	<i>FSTS</i>	.9137	.8488
	<i>FATA</i>	.8808	.7125
	<i>PDIO</i>	.7465	.7125
	<i>TMIE</i>	.6806	.4632
Eigenvalue 2.8133			

The final test was assessing the normality of the distribution of the values of the seventy-four companies on the revealed scale of DOI_{INTS} . Doing so involved treating *FSTS*, *FATA*, *OSTS*, *TMIE*, and *PDIO* as a linear combination and summing the value per attribute for each MNC (see Table 5). Because the number of classes significantly influences the contour of the distribution, we determined class width as advised by Anderson, Sweeny and Williams [1983]. The plot of firms' value for DOI_{INTS} suggested a normal distribution. Test of the goodness of fit of the distribution confirmed this inference: the computed X^2 of 5.73 was less than the test statistic of $X^2_{.05} = 5.99$. The plot of the standardized scale scores also satisfied Chebyshev's Theorem.

Assessing the degree of criterion bias due to the use of *Forbes* data prompted two tests. The first considered the degree of correspondence among the rankings of the seventy-four MNCs on DOI_{INTS} , *FSTS*, *FATA*, *OSTS*, *PDIO*, and *TMIE*. The frequency of correspondence existed to the degree one would expect by chance.⁸ The second test considered Kendall's *tau* between DOI_{INTS} and each of the five single measures. All associations were significant at the .0001 level.⁹

TABLE 5
Sampled Company's Value on the Measurement Scale of DOI_{INTS}

Company	DOI_{INTS}	Company	DOI_{INTS}
1. CPC	3.13	38. Albbot	2.10
2. Exxon	3.02	39. Allied Signal	2.10
3. Colgate-Palmolive	2.97	40. Caterpillar	2.08
4. IBM	2.91	41. Kimberly Clark	2.08
5. Gillette	2.90	42. Borden	2.00
6. Mobil	2.86	43. Philip Morris	2.00
7. Pfizer	2.83	44. Alcoa	1.97
8. American Brands	2.82	45. Dresser	1.95
9. DEC	2.81	46. Emerson Electric	1.94
10. Hewlett Packard	2.77	47. Union Carbide	1.94
11. Dow Chemical	2.75	48. Amoco	1.92
12. Merck	2.65	49. PPG	1.91
13. Johnson&Johnson	2.62	50. Quaker Oats	1.90
14. Coca-Cola	2.61	51. General Motors	1.88
15. Eli Lilly	2.59	52. Scott Paper	1.82
16. AMP	2.58	53. Rockwell Int'l	1.81
17. Avon	2.50	54. WR Grace	1.81
18. Motorola	2.45	55. Deere	1.75
19. Warner Lambert	2.45	56. Tenneco	1.74
20. Black & Decker	2.44	57. TRW	1.64
21. NCR	2.43	58. Baxter Travenol	1.64
22. Texaco	2.35	59. Sara Lee	1.63
23. 3M	2.33	60. Pepsi Cola	1.57
24. Procter & Gamble	2.32	61. Ralston-Purina	1.57
25. American Cyanamid	2.31	62. International Paper	1.54
26. Goodyear	2.31	63. United Technologies	1.52
27. Unisys	2.30	64. Chevron	1.51
28. Heinz	2.29	65. Honeywell	1.44
29. Bristol-Myers	2.29	66. Xerox	1.38
30. Kellogg	2.27	67. Whirlpool	1.38
31. Eastman-Kodak	2.25	68. Atlantic Richfield	1.18
32. Intel	2.23	69. Westinghouse	1.13
33. Monsanto	2.17	70. General Electric	1.06
34. Ford	2.14	71. Phillips Petroleum	1.00
35. E.I. Dupont	2.14	72. Chrysler	.94
36. Amer Home Products	2.13	73. Unocal	.91
37. Texas Instruments	2.12	74. Sun	.54

Method of Calculation

The five variables constituting DOI_{INTS} —*FSTS*, *FATA*, *OSTS*, *PDIO*, *TMIE*—are ratio variables. As such, the range of value for a firm is 0.0 (absolutely no international involvement) to 5.0. (absolutely total international involvement). We calculate a firm's score through the following operation:

$$FSTS + FATA + OSTS + PDIO + TMIE = DOI_{INTS}$$

So, for example, IBM's value of 2.91 for DOI_{INTS} was derived accordingly:

$$.59_{FSTS} + .50_{FATA} + .77_{OSTS} + .9_{PDIO} + .15_{TMIE} = 2.91.$$

TABLE 6
Company Rankings on Six Estimators of the Degree
of Internationalization of a Firm

Company	DOI_{INTS}	FSTS	FATA	OSTS	TMIE	PDIO
CPC	1	12	9	14	1	1
Exxon	2	1	1	26	26	28
Colgate Palmolive	3	3	3	28	19	20
IBM	4	6	10	8	21	15
Gillette	5	2	2	19	10	46
Mobil	6	4	8	49	7	1
Pfizer	7	22	31	4	2	18
American Brands	6	15	11	1	38	17
DEC	9	9	7	3	25	43
Hewlett Packard	10	12	20	5	29	19
Dow Chemical	11	8	5	31	16	14
Merck	12	17	30	9	13	17
Johnson & Johnson	13	14	14	11	70	21
Coca-Cola	14	6	4	35	31	39
Eli Lilly	15	33	25	2	23	8
AMP	16	10	16	7	57	38
Avon	17	8	6	38	14	48
Motorola	18	13	29	13	43	41
Warner Lambert	19	18	40	12	5	57
Black & Decker	20	23	13	37	15	40
NCR	21	4	17	15	35	62
Texaco	22	20	64	51	6	1
3M	23	21	25	46	9	35
Procter & Gamble	24	39	51	23	3	34
American Cyanamid	25	28	23	33	22	37

DISCUSSION

Statistically, the results suggest that DOI_{INTS} provides a meaningful measure of the DOI of a firm. The significance of the item-scale correlations, satisfactory *alpha* coefficient,¹⁰ parsimonious factor structure, and normality of the distribution confirmed that DOI_{INTS} met statistical standards. The exclusive use of archival data to operationalize the attribute variables, by enabling replication, encourages developing a functional standard of measurement. Conceptually, we believe DOI_{INTS} is an improvement over other measurement methods. In the least, the chronic arbitrariness of measures of DOI suggests that any progress is useful. More fundamentally, DOI_{INTS} begins to address the threat of measurement error to the validity of research findings by providing a method to disentangle the misleading influences of these errors before theory testing. Because DOI_{INTS} integrates the instrumental and positivistic methods, it should benefit from the advantages of both techniques while diminishing their deficiencies. Nunnally [1978:264] observes that "regardless of what is found in item analysis, the final decision to include or reject an item is based primarily on human judgement." The results of our study

make such deliberation incidental. The elements of DOI_{INTS} —performance (*FSTS*), structural (*FATA* and *OSTS*), and attitudinal (*TMIE* and *PDIO*)—conform with theory.

A Case in Point Revisited

Recall our earlier assertion that an explanation for the disarray of Table 1 was the seemingly inevitable use of *FSTS* as the estimator of *DOI*. A test of this allegation is in order. In principle, Anderson, Sweeny and Williams [1983] explain that the goodness of the inferential process is a function of the goodness of fit of the distribution. In practice, presupposing that we had used *FSTS* as the estimator of *DOI* in evaluating some attribute of the same seventy-four MNCs, would the distribution of *FSTS* be normal? If yes, then this outcome would challenge the alleged superiority of a DOI_{INTS} versus *FSTS*. Conversely, the non-normality of the distribution of *FSTS* would testify to its tendency toward excessive measurement error.

To preempt concerns of rigor, the data used for this test were the three-year average, 1988-1990, of *FSTS* for the seventy-four MNCs. Such short-term smoothing helps neutralize irregularities in the data and improves the odds of a normal distribution. The plot of the distribution of *FSTS* exhibited irregularities. The computed *chi*-square of the distribution's goodness of fit, $\chi^2_{.05}=25.62$, exceeds the test statistic, $\chi^2_{.05}=5.99$.

This result supports the principle that use of a single-item estimator of a construct creates the risk that any unusual circumstances that might distort the normal validity of the measure will contaminate or ruin the goodness of the results. More precisely, a research design that uses *FSTS* as the sole indicator of *DOI* is vulnerable to the risk that a firm's foreign sales in the period of study may be artificially inflated or deflated by some conceptually irrelevant factor having nothing to do with the "true" internationalization of a firm. For example, because MNCs consolidate foreign sales from a variety of national currencies into a single numeraire, extreme variance in the portfolio of currencies across firms or simply a random shock in currency rates can ruin the meaningfulness of *FSTS*. Since the nature of a single measure precludes estimating its reliability, the use of such measures as the basis for empirical study unwittingly creates the potential for biased data, flawed analyses, and trivial or misleading results.

Lastly, this comparison of single- and multi-item measures corroborates this paper's theme that a multi-item scale, such as DOI_{INTS} , is more likely to tap a broader range of the valid content of the total meaning of the construct of a firm's *DOI*. As such, the use of a multi-item scale goes further to ensure that circumstances that may invalidate one dimension of the measurement scale are not so apt to invalidate the entire measurement scale. If that error is endemic to measurement, then the use of a multi-item scale enables errors to cancel out or exert more of a random effect, thereby improving the goodness of the data, and ultimately, the goodness of inferences drawn from the data.

CONCLUSION

The psychometric literature asserts that reliably discriminating the domain and attributes of a construct depends on minimizing measurement error. We believe measuring DOI with the linear combination of *FSTS*, *FATA*, *OSTS*, *PDIO*, and *TMIE* does so by reducing the error that results from sample, systematic, and random bias. Coombs [1964] points out that a study's sample is a major source of error. Notwithstanding variation in their absolute and relative internationalization, scholars typically treat "MNCs" as isomorphic in their sampling method. The intrinsic limitations of positivistic and instrumental approaches compound sampling error by aggravating existing systematic error—i.e., consistently over-reporting or under-reporting the attributes of the sample—and random error—i.e., inadvertently reinforcing or diluting the observed relationships among variables. If unchecked, these sources of measurement error distort estimates of the magnitude and relevance of the actual relationship and preclude distinguishing trait variance from unwanted method variance, thereby inducing errors in inference [Bagozzi, Youjae and Phillips 1991; Cook and Campbell 1979]. Our results lead us to conclude that DOI_{INTS} , by providing a method to improve the goodness of the sample as well as clarifying the extent of random and systematic error, fortifies the reliability of measurement and validity of interpretation.

Ultimately, the challenge of this area of study is transforming the diversity of empirical reports on the internationalization of a firm into received theory. Table 2 suggests that theory testing remains ambiguous precisely because we cannot ascertain whether the acceptance or rejection of a hypothesis is the result of excessive error in measurement or the adequacy or inadequacy of prevailing theories of the internationalization of the firm. Even if an intuitively useful relationship is detected, such as Johanson and Vahlne's [1977] notion that psychic distance moderates the incremental path of the international expansion of a firm, the magnitude of the relationship can be significantly understated or overstated because of measurement error. This, in turn, raises questions regarding the explanatory power of the model, as exemplified by the subsequent reports of Millington and Bayliss [1990] and Sullivan and Bauerschmidt [1990] that reject the alleged influence of psychic distance. If unchecked, such uncertainty reduces the practical relevance of research results.

Schoenfeldt [1984:78] notes that "The construction of the measuring device is perhaps the most important segment of any study. Many well-conceived research studies have never seen the light of day because of flawed measures." Korman [1974] and Schwab [1980] echo these concerns, concluding that theoretical progress is simply not possible without adequate measurement. We have argued that progress demands a coherent approach to determine the reliability of measurement and the validity of findings. We believe that the measurement scale of DOI_{INTS} contributes a meaningful method to facilitate progress.

FUTURE RESEARCH

While content validity falls within the scope of a single study, establishing construct validity requires evidence from many studies using different measures [Cook and Campbell 1979; Fiske 1982; Nunnally 1978]. The construct validity of DOI can only be inferred through the pattern and magnitude of covariation among multiple measures of the construct and a comparison of these measures to measures of one or more other constructs. Conducting this analysis requires estimating discriminant and convergent validity. Estimating each property requires the input of multiple as well as dissimilar methods and measures. Such estimation is not currently possible. Granted, instrumental and positivistic research suggest a few potential measures. Still, we lack the multiplicity of dissimilar, empirical measures of DOI that would provide the inputs into the procedures of construct validation. More precisely, methods of construct validation, such as multitrait-multimethod, first-order confirmatory factor analysis, second-order confirmatory factor analysis, and hierarchical confirmatory factor analysis, impose stringent methodological standards. Each technique demands multiple, dissimilar, and tested empirical measures of the construct that, at the least, have exhibited some degree of content validity. Currently, the research reports of international business, with regard to providing a diversity of empirical measures and methods of estimating DOI with more than simply face validity, do not permit satisfying the general, to say nothing of the more specific, conditions of construct validation procedures.

Future research can help build the inventory of multiple, dissimilar measures and methods required to evaluate the construct validity of DOI. The limitations of our design suggest two immediate points of innovation for these efforts. First, other measures may complement or replace the items that comprise *DOI_{INTS}*. Try as we did, we found this set was not exhaustive. Imagination may lead to more creative uses of archival data. Second, testing the reliability of our scale or another scale in different populations promises to refine the reliability of measuring DOI. Granted, Chandler [1986] confirms the relevance of sampling the traditional domain of large, American MNCs. Still, our sample does not represent small- or medium-sized exporters and MNCs, service companies, or non-American firms.

NOTES

1. The sample of 74 MNCs includes 49 firms that have been listed every year plus the 25 firms that are also listed in 1990 but that are not members of the group of 49. Regarding the latter, these firms have been ranked anywhere from 1 to 10 years, with the distribution skewed toward higher frequencies. The 46 manufacturing firms excluded from the study were ranked early or sporadically, or have since gone private, merged, or been acquired.

2. Chandler [1986:444] reported that American MNCs "are still clustered in much the same type of industries, and in most of these industries the leaders in the first decade of the century were still the leaders in the ninth decade."

3. This statistic may underestimate Top Managers' International Experience given the possibility of career changes and unreported company affiliations. However, we assume that this variation is systematic

across the sample given reporting requirements imposed by the *Directory* as well as protocols governing the format and breadth of professional resumes.

4. Johanson and Vahlne [1977] conclude that a firm's international expansion radiates from the home market in a systematic fashion: the operations of the firm are extended first to those markets that best fit the cognitive and resource character of the company, and ultimately those of the poorest fit. A quasi-concentric expansion pattern emerges because extending the firm's international operations requires managing the progressively complex contingencies of progressively less similar markets.

5. Ronen and Shenkar's [1985] meta-analysis of cross-cultural studies decomposed the world into ten psychological zones—"Anglo, Germanic, Nordic, Near Eastern, Arab, Far Eastern, Latin American, Latin European, Independent, and Other." We used these categories as a template to assess the psychic distribution of each firm's subsidiaries. For instance, if Company Y reported the following distribution of units,

National Site of Subsidiary	Psychic Zone
1. Austria	1. Germanic
2. Canada	2. Anglo
3. Belgium	3. Latin European
4. Chile	4. Latin American
5. Hong Kong	5. Far Eastern
6. Greece	6. Near Eastern
7. Bermuda	7. Other

then it was credited with operating a subsidiary in the corresponding zone, and assigned a score of 70%.

6. One can use any of three techniques to establish the reliability of measurement: internal consistency is concerned with the degree to which the items on an instrument are interrelated; test-retest directs attention to the degree to which a scale is stable over time; and, interrater reliability is concerned with the degree of agreement among multiple independent judges. The inappropriateness of the test-retest and interrater methods to our research question led to evaluating the internal consistency of a scale of single items.

7. Rumelt [1974] suggests that the relationship between the "type" of firm, as defined by the degree of "relatedness" among the firm's products, discriminates the character of industrial companies. The degree of relatedness among the firm's products, Rumelt notes, fundamentally influences product life cycle pressures and the potential for synergy. As such, different types of firms experience different levels of product pressure to internationalize operations.

8. With regards to the intersection in the rankings of firms in terms of DOI_{INTS} and $FSTS$, only two firms share the same rank. For DOI_{INTS} and $FATA$, four firms share the same rank; for DOI_{INTS} and $OSTS$, six firms share the same rank; for DOI_{INTS} and $TMIE$, five firms share the same rank; and for DOI_{INTS} and $PDIO$, four firms share the same rank.

9. The strength of each association was such: DOI_{INTS} and $FSTS$ ($\tau = .68$), DOI_{INTS} and $FATA$ ($\tau = .64$), DOI_{INTS} and $OSTS$ ($\tau = .57$), DOI_{INTS} and $TMIE$ ($\tau = .41$), DOI_{INTS} and $PDIO$ ($\tau = .45$).

10. Nunnally [1978] notes that an α of .70 is the lower bound for an exploratory effort. Past this threshold, one can be reasonably certain that there is strong covariance among the items as well as that the sampling domain has been adequately captured.

REFERENCES

- Adler, Nancy, Robert Doktor & S. Gordon Redding. 1986. From the Atlantic to the Pacific century. *Journal of Management*, 12(2): 295-318.
- Anderson, David, Dennis Sweeny & Thomas Williams. 1983. *Quantitative methods for business*. New York: Praeger.
- Bagozzi, Richard P., Yi Youjae & Lynn Phillips. 1991. Assessing validity in organizational research. *Administrative Science Quarterly*, 36: 421-58.
- Beamish, Paul W. & Richard C. daCosta. 1984. Factors affecting the comparative performance of multinational enterprises. *EIBA 1984 Conference*. Rotterdam: Erasmus University.
- Bollen, Kenneth. 1989. *Structural equations with latent variables*. New York: Basic Books.
- Buckley, Peter J. & Mark Casson. 1985. *The economic theory of the multinational enterprise*. New York: St. Martin's Press.

- Buckley, Peter J., John H. Dunning & Robert B. Pearce. 1977. The influence of firm size, industry, nationality, and degree of multinationality in the growth and profitability of the world's largest firms. *Weltwirtschaftliches Archiv*, 114: 243-57.
- Buhner, Rolf. 1987. Assessing international diversification of West German corporations. *Strategic Management Journal*, 8: 25-37.
- Campbell, Donald & Donald Fiske. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56: 81-105.
- Capon, Noel, John U. Farley & Scott Hoeting. 1990. Determinants of financial performance. *Management Science*, 36(10): 1143-59.
- Caves, Richard E. 1982. *Multinational enterprise and economic analysis*. Cambridge, U.K.: Cambridge University Press.
- Cavusgil, S. Tamer & Yezdi M. Godiwalla. 1982. Decision-making for international marketing. *Management Decision*, 20: 48-57.
- Chandler, Alfred. 1986. The evolution of modern global competition. In Michael Porter, editor, *Competition in global industries*. Boston: Harvard Business School Press.
- Collins, J. Markham. 1990. A market performance comparison of U.S. firms active in domestic, developed and developing countries. *Journal of International Business Studies*, 2: 271-87.
- Cook, Thomas D. & Donald T. Campbell. 1979. *Quasi-experimentation: Design and analysis issues in field settings*. Boston: Houghton Mifflin.
- Coombs, Charles. 1964. *A theory of data*. New York: Wiley.
- Daniels, John D. & John Bracker. 1989. Profit performance: Do foreign operations make a difference? *Management International Review*, 29(1): 46-56.
- Eppink, D. Jan & B. Van Rhijin. 1988. The internationalization of Dutch industry. *Long Range Planning*, 21(5): 54-60.
- Ericsson, Anders & Herbert A. Simon. 1980. Verbal reports as data. *Psychological Review*, 2: 215-51.
- Fiske, Donald. 1982. Convergent-discriminant validation in measurements and research strategies. In D. Brinberg & L. Kidder, editors, *Forms of validity in research*, 77-92. San Francisco: Jossey-Bass.
- Forsgren, Mats. 1989. *Managing the internationalization process: The Swedish case*. London, U.K.: Routledge.
- Franko, Lawrence. 1989. Global corporate competition: Who's winning, who's losing, and the R&D factor as one reason why. *Strategic Management Journal*, 10(2): 49-74.
- Geringer, J. Michael, Paul W. Beamish & Richard C. daCosta. 1989. Diversification strategy and internationalization: Implications for MNE performance. *Strategic Management Journal*, 10(2): 109-19.
- Grant, Robert M. 1987. Multinationality and performance among British manufacturing companies. *Journal of International Business Studies*, 18(3): 79-89.
- _____, Azar P. Jammine & Howard Thomas. 1988. Diversity, diversification, and profitability among British manufacturing companies, 1972-1984. *Academy of Management Journal*, 31(4): 771-801.
- Hofstede, Geert. 1993. Cultural constraints in management theories. *Academy of Management Executive*, 7(1): 81-94.
- Horst, Thomas E. 1973. Firm and industry determinants of the decision to invest abroad. *Review of Economics and Statistics*, 54(August): 258-66.
- Johanson, Jan & Jan-Erik Vahlne. 1977. The internationalization process of the firm. *Journal of International Business Studies*, 8: 23-32.
- Keown, Charles F., N. Synodinos, L. Jacobs & Reginald Worthley. 1989. Transnational advertising-to-sales ratios: Do they follow the rules? *International Journal of Advertising Age*, 8(4): 375-82.
- Korman, Abraham. 1974. Contingency approaches to leadership. In J. Hunt & L. Larson, editors, *Contingency approaches to leadership*. Carbondale, Ill.: Southern Illinois University Press.
- Kumar, Manmohan S. 1984. *Growth acquisition and investment: An analysis of the growth of industrial firms and their overseas activities*. Cambridge, U.K.: Cambridge University Press.
- Maisonrouge, Jacques G. 1983. Education of a modern international manager. *Journal of International Business Studies*, 13: 56-60.
- March, James G. & Herbert A. Simon. 1958. *Organizations*. New York: John Wiley.

- McDonald, Richard P. & Herbert W. Marsh. 1990. Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin*, 107: 247-56.
- Michel, Allen & Israel Shaked. 1986. Multinational corporations vs. domestic corporations: Financial performance and characteristics. *Journal of International Business Studies*, 18(3): 89-100.
- Millington, Andrew I. & Brian T. Bayliss. 1990. The process of internationalization: UK companies in the EC. *Management International Review*, 30(2): 151-61.
- Mitchell, Terence. 1985. An evaluation of the validity of correlational research conducted in organizations. *Academy of Management Review*, 10(2): 192-206.
- Nunnally, Jum C. 1978 (second edition). *Psychometric theory*. New York: McGraw-Hill.
- Perlmutter, Howard V. 1969. The tortuous evolution of the multinational corporation. *Columbia Journal of World Business*, 4 (January-February): 9-18.
- Phillips, Lynn W. & Richard P. Bagozzi. 1985. On measuring organizational properties of distribution channels: Methodological issues in the use of key informants. *Research in Marketing*, 8: 313-69.
- Ronen, Simcha & Oded Shenkar. 1985. Clustering countries on attitudinal dimensions: A review and synthesis. *Academy of Management Review*, 10(3): 435-54.
- Rugman, Alan M., Donald J. Lecraw & Laurence D. Booth. 1985. *International business: Firm and environment*. New York: McGraw-Hill.
- Rumelt, Richard P. 1974. *Strategy, structure, and performance*. Boston: Harvard Business School Press.
- Schmidt, Frank L. & John E. Hunter. 1981. Employment testing: Old theories and new research findings. *American Psychologist*, 36: 1128-37.
- Schoenfeldt, Lyle F. 1984. Psychometric properties of organizational research instruments. In T. S. Bateman & G. R. Ferris, editors, *Method and analysis in organizational research*. Reston, Va.: Reston Publishing Company.
- Schwab, Donald. 1980. Construct validity in organizations behavior. In B. Staw & L. Cummings, editors, *Research in organizational behavior*, Vol. 2, 3-43. Greenwich, Conn.: JAI Press.
- Seidler, L. John. 1974. On using informants: A technique for collecting qualitative data and controlling for measurement error in organizational analysis. *American Sociological Review*, 39: 816-31.
- Shaked, Israel. 1986. Are multinational corporations safer? *Journal of International Business Studies*, 17(1): 75-80.
- Siddharthan, Natteri & Sanjaya Lall. 1982. Recent growth of the largest U.S. multinationals. *Oxford Bulletin of Economics and Statistics*, 44 (February): 1-13.
- Stopford, John M. & John H. Dunning. 1983. *The world directory of the multinational enterprises 1982-83*. Detroit, Mich.: Gale Research Company.
- Stopford, John M. & Louis T. Wells. 1972. *Managing the multinational enterprise*. New York: Basic Books.
- Sullivan, Daniel. 1993. The "threshold of internationalization": Replication, extension, and reconsideration. *Management International Review*, in press.
- _____. & Alan Bauerschmidt. 1989. Common factors underlying barriers to export: A comparative study in the European and U.S. paper industry. *Management International Review*, 29(2): 46-63.
- _____. 1990. Incremental internationalization: A test of Johanson and Vahlne's thesis. *Management International Review*, 30(1): 19-30.
- Turnbull, Peter W. 1987. A challenge to the stages' theory of the internationalization process. In Philip J. Rosson & Stan D. Reid, editors, *Managing export entry and expansion*. New York: Praeger Publishers.
- Vernon, Raymond. 1971. *Sovereignty at bay: The multinational spread of U.S. enterprises*. New York: Basic Books.
- Welch, Lawrence & Reijo Luostarinen. 1988. Internationalization: Evolution of a concept. *Journal of General Management*, 14(2): 34-55.
- Yoshihara, Hiro. 1985. Multinational growth of Japanese manufacturing enterprises in the postwar period. *Proceedings of the Fuji International Conference on Business History*. Tokyo: University of Tokyo Press.