Does Diversity Trump Ability?

An Example of the Misuse of Mathematics in the Social Sciences

Abigail Thompson

What Is the Issue and Why Should We Care?

"Diversity" has become an important concept in the modern university, affecting admissions, faculty hiring, and administrative appointments. In the paper "Groups of diverse problem solvers can outperform groups of high-ability problem solvers" [1], L. Hong and S. Page claim to prove that "To put it succinctly, diversity trumps ability." We show that their arguments are fundamentally flawed.

Why should mathematicians care? Mathematicians have a responsibility to ensure that mathematics is not misused. The highly specialized language of mathematics can be used to obscure rather than reveal truth. It is easy to cross the line between analysis and advocacy when strongly held beliefs are in play. Attempts to find a mathematical justification for "diversity" as practiced in universities provide an instructive example of where that line has been crossed.

In this article we examine the arguments of the Hong and Page paper in detail. The paper contains what the authors call a "Mathematical Theorem," ostensibly proving that a group picked on the basis of "diversity" criteria outperforms one picked on the basis of "ability." In contrast to much of the diversity research literature, this paper claims to be based on mathematical reasoning. Its publication in 2004 in the *Proceedings of the National Academy of Sciences* has given it credibility, and it is widely cited. Its conclusions are presented as mathematical truth. Referring to this work in his 2007 book *The Difference* [3] (p. 165), Page says,

> …the veracity of the diversity trumps ability claim is not a matter of dispute. It's true, just as 1 + 1 = 2 is true.

Under careful scrutiny, however, the paper is seen to have essential and irreparable errors.

The mathematical content of [1] is presented in two main sections. In the first of these, "A Computational Experiment," the authors describe a computer simulation involving a collection of algorithms working together to solve a simple optimization problem. In this section the authors find that one collection of algorithms outperforms a second collection. They assign the label "diversity" to the first collection and the label "ability" to the second, and conclude that this is evidence that "diversity trumps ability." In a subsequent section titled "A Mathematical Theorem," the authors indicate that their analysis "…explores the logic behind the simulation results and provides conditions under which diversity trumps ability."

There are multiple problems in each of these sections. We can summarize the content of the theorem as follows: suppose that a group of people is set a task and the entire group's performance is compared to that of just one member of the group working alone on the same task. Assume also that the conditions of the task are such that one person working alone can never complete the task, and that the whole group working together will always complete the task. It is neither surprising nor difficult to see that, in this situation, the group as a whole will outperform the individual. Yet this is the entire content of what Hong and Page call the "Mathematical Theorem," Theorem 1 in [1].

Abigail Thompson is professor of mathematics at the University of California, Davis. Her email address is `thompson@math.ucdavis.edu`.

DOI: http://dx.doi.org/10.1090/noti1163

We also point out several issues with the "Computational Experiment" section of [1]. Here the authors discuss computer simulations intended to illustrate and support the conclusions of Theorem 1. We demonstrate that Theorem 1 is unrelated to the computational experiment and that the experiment offers no support for the social applications proposed by the authors.

A Mathematical Theorem

We first consider the section "A Mathematical Theorem" [1]. Theorem 1 of that section is the basis of the claim that there is a mathematical proof that diversity trumps ability. Once the unnecessary technicalities are removed and basic errors corrected, the theorem is revealed to be little more than a straightforward restatement of its hypotheses. Furthermore, a careful examination of Theorem 1's statement shows that it has no real-world applications.

Statement of the Theorem

Theorem 1 concerns the problem of finding the maximum value of a fixed real-valued function V defined on a finite set X. The function V is assumed to attain its maximum at a unique point x^* in X. Attempting to find this maximum is a finite collection of algorithms, which Hong and Page call "agents," or "problem-solvers." The collection of k agents (algorithms) is denoted by Φ. An agent ϕ is a function from X to X such that $V(\phi(x)) \geq V(x)$. Depending on the initial point x in X, an agent ϕ will sometimes but not always return the point x^* at which V achieves its maximum. Agents can work together in some way on the problem of finding the maximum of V. The authors assume that two copies of a single agent working together operate sequentially using composition of functions, with one taking as input the output of another. Combined with their definition of an agent, this implies that multiple copies of one agent perform identically to a single copy, an assumption required for their proof. We note that this is a strikingly restrictive and artificial condition, precluding an intelligent division of work.

Each agent produces an average value for V, by averaging over all starting points in X with equal weight. Agents can be ordered by these average values, and one agent is said to be *better* than another if its average is larger.

Hong and Page then make the following assumptions.

- Assumption 1: $\forall \phi \in \Phi, \exists x \in X$ such that $V(\phi(x)) < V(x^*)$; i.e., for each agent, there is some starting point for which V of its stopping point is not the global maximum.
- Assumption 2: $\forall x \in X$, $x \neq x^*$, $\exists \phi \in \Phi$ such that $\phi(x) \neq x$; i.e., no point of X is fixed under all elements of Φ except x^*.
- Assumption 3: Φ has a unique best element.

Additionally, as part of the definition of an agent, they include:

- Assumption 0: (i) $\forall x \in X, V(\phi(x)) \geq V(x)$. (ii) $\phi(\phi(x)) = \phi(x)$.

Hong and Page offer this interpretation of Assumption 2: "When one agent gets stuck, there is always another agent that can find an improvement" ([1], p. 16387). This interpretation is incorrect without the additional hypothesis that $V(x)$ is a one-to-one function. This error gives rise to a counterexample to the theorem, described in the Appendix. We proceed with the additional assumption that V is one-to-one.

Given this additional hypothesis, together with the additional assumption that agents "work together" by successive composition of functions, Hong and Page's assumptions imply:

(1) An agent working alone will sometimes not return the point x^*.
(2) All agents working together will always return the point x^*.
(3) There is a unique best agent.
(4) Multiple copies of a given agent working together perform identically to a single copy.

From (2) we see that the complete collection of all k agents in Φ, working together, will always return the point where the maximum value of the function V occurs, irrespective of the initial starting point. In contrast, k copies of the best agent in Φ behave identically to a single copy of the best agent (4) and thus do not always return the point where the maximum value of the function V occurs (1). Thus the complete collection of all k agents in Φ, working together, performs better than k copies of the best agent in Φ. Theorem 1 amounts to little more than this simple observation. The following statement of the theorem, as given in ([1], p. 16388), may sound more impressive.

Theorem 0.1 (Theorem 1)**.** *Let Φ be a group of problem solvers that satisfy assumptions 1-3. Let μ be a probability distribution over Φ with full support. Then, with probability one, a sample path will have the following property: there exist positive integers N and N_1, $N \geq N_1$, such that the joint performance of the N_1 independently drawn problem solvers exceeds the joint performance of the N_1 individually best problem solvers among the group of N agents independently drawn from Φ according to μ.*

In Page's book *The Difference* this theorem has been named the "Diversity Trumps Ability" theorem ([3], p. 162), and Page offers this application:

> How do we apply this in the real world? Simple. When picking two hundred employees from a pool of thousands, provided the people are all smart, we should keep the theorem in mind and not necessarily rank people by some crude ability score and pick the best. We should seek out difference.

While it may sound somewhat like Theorem 1, this interpretation is not correct. In reality Theorem 1 has nothing to say about hiring employees. The principal reason it does not apply involves the somewhat mysterious presence in the formal statement of Theorem 1 of the numbers N and N_1. What is not clearly stated is that N and N_1 can, and generally will, be substantially larger than k, the size of the initial pool of agents (or employees). To apply Theorem 1 to pick employees, you must be willing and able to make large numbers of clones of each of your job applicants, and you must be interested in picking from this army of clones a staff of tens of thousands, or the theorem has nothing to say about your hiring process.

We now examine in more detail the authors' arguments as they take a detour through probability. Our goal is to clarify the statement and proof of Theorem 1, and the nature (and relative magnitude) of the unspecified numbers N and N_1.

Idea of the Proof

We illustrate the proof of the corrected theorem with a simple example, using the case $k = 6$. In this case there are six distinct agents in Φ. We can dispense with much of the technical language from probability by associating each of these agents to a face of a standard die, to facilitate picking them at random. (While we confine ourselves to the case $k = 6$ and equal probabilities for purposes of clarity, the arguments apply more generally. For arbitrary finite k we can use any probability distribution with full support on k agents to select agents at random.) Of critical importance, we need to assume there are an unlimited number of copies of each agent, so that, for example, the sentence "Pick fifty agents at random from among the six agents" makes sense. In contrast, in the process of selecting job applicants in the real world, a request to "Pick fifty workers to hire at random from among six job applicants" does not make sense. Suppose that Agent 1 is the unique best agent of our six.

We are ready to understand the essential argument of Theorem 1 in three steps:

Step 1: Throw the die fifty times, and record the results. Call the corresponding collection of fifty agents "Group A." With high probability, Group A contains at least one copy of each of the six agents. Of course Group A probably contains several copies of each agent, but that's okay; we just want to make the size of the group large enough to be reasonably certain that it will contain at least one copy of each.

Step 2: Now throw the die 10,000 times, and record the results. It is extremely likely that each face of the die will show up at least fifty times in the results. In particular, with high probability Agent 1 will show up at least fifty times in the results. Hence if we select the best fifty agents from among the 10,000 with high probability we will select fifty copies of Agent 1. Call this collection of the best fifty agents "Group B."

Step 3: Since with high probability Group A includes a copy of each of our original six agents, Group A will, with high probability, always find the point where the maximum value of V occurs (2). With high probability Group B, however, is fifty copies of Agent 1 and this group of fifty will not always find the point where the maximum value of V occurs ((1) and (4)). Conclude that Group A outperforms Group B.

In this example the numbers N and N_1 have values $N = 10,000$ and $N_1 = 50$. The conclusion, stated in English, would read something like this:

Given six distinct problem-solvers, if fifty are selected at random from among these six, they will, with high probability, collectively outperform the fifty best problem-solvers chosen from 10,000 selected at random from among the six.

This is, as we have shown, an easy consequence. Notice that, when stated this way, it does not sound very sensible. One does not, in general, talk about selecting fifty "problem-solvers" from a group of six. This example highlights a misuse of the word "problem-solvers" in the formal statement of Theorem 1. A "problem-solver" strongly suggests an individual person. However as Hong and Page are using the word, a "problem-solver" is an algorithm. Algorithms, unlike people, can be made to duplicate each other exactly. Set ten copies of a single algorithm to painting a house, and they will paint the same wall ten times over. Ten humans are unlikely to do so.

The problem revealed in the case $k = 6$ does not disappear when using a larger initial set of distinct problem-solvers, or agents. Regardless of the size k of the initial pool of distinct problem-solvers, the argument remains very much the same. There is no information given by the theorem about the performance of any proper subset of the initial pool. The numbers N and N_1 not only may be much larger than k, but indeed must be so for the proof to work. The calculation of appropriate values of N and N_1 is an instance of the classic "coupon collector's problem" from standard probability theory ([2], p. 32). The passage from "highly likely" to certainty, as claimed in Theorem 1, requires a consideration of what happens in the limit as N_1 and N go to infinity.

What can reasonably be concluded from the outperformance of Group A? Nothing. We should not be even mildly surprised to find that a group which includes the best agent along with a collection of additional agents outperforms a group consisting only of identical copies of the best agent.

Furthermore, there is a curious and crucial discrepancy between the mathematical argument and the "diversity trumps ability" terminology. Somehow Hong and Page have transformed Group A, whose chief advantage is that it contains a copy of every single agent in the pool, into a "diverse" group. They say, "This result relies on the intuition that, as the initial pool of problem solvers becomes large, the best-performing agents necessarily become similar in the space of problem solvers. Their relatively greater ability is more than offset by their lack of problem-solving diversity" ([1], p. 16385). This claim doesn't appear to have any mathematical meaning; the term "diversity" has not been defined in the context of Theorem 1.

Having disposed of the ideas that Theorem 1 either contains substantial mathematical content, or is somehow applicable to real life, we now turn to the section of [1] containing a computational simulation.

A Computational Experiment

We give a description of the "Computational Experiment" described in [1] (p. 16386). Let $X = [1, 2, 3, \ldots, n]$ be the set consisting of the first n integers, and let V be a function from X to the real numbers. The goal posited in the computational model in [1] is to find the maximum value of V on the set X. The function V is assumed to have a unique maximum at x^*.

Fix two integers, l and k, with $1 \leq k < l < n$, and define an *agent* to be a list of k distinct integers in $[1, \ldots, l]$. An agent describes a procedure to find a maximum value of V as follows:

The agent $\alpha = (a_1, a_2, a_3, \ldots, a_k)$ starts at some point i of X. It checks the value of V at i and then at $i + a_1$. If $V(i) \geq V(i + a_1)$, α next checks the value of V at $i + a_2$. If $V(i) < V(i + a_1)$, α next checks the value of V at $i + a_1 + a_2$. The search continues for elements of X (mod n) and for successive integers in α (mod k) until α gets stuck for a full k checks. Call this the "stopping point" of α for i, and denote it by $\alpha(i)$. Thus each α is a function from X to X. The value of V at the stopping point i is a *local optimum* for α. It is uniquely determined by α and the starting point i. We note in passing that this is an odd and inefficient way to go about finding the maximum value of a function on a finite set.

An agent α has an *average value* on X, defined by $(1/n)\sum_{i=1}^{n} V(\alpha(i))$. An agent α is said to be *better* than an agent β if α has a higher average value than β.

Hong and Page describe a simulation with $n = 2000$, $l = 20$, and $k = 3$. This gives a pool of $20 \times 19 \times 18 = 6840$ agents. They select ten at random and the ten best from the entire pool, and compare their performance as groups. They are surprised to find that the ten random agents acting together outperform the ten best agents acting together. We note that Theorem 1 offers no insight into this experiment, since the assumptions of the theorem are not met by the set-up of the experiment.

To understand what this simulation is doing, we ran a computer simulation following the description in [1] and were able to reproduce the results (code available by request). To see why the result of this simulation is not surprising, we take a closer look at the data from one run of the program. Since Hong and Page's simulation was based on an unreported random function on 2000 points, we used a function on 2000 points constructed with a random number generator.

For our simulation, we obtained the following list of the ten best agents:

$$[12, 4, 13], [7, 9, 14], [4, 12, 13], [10, 6, 17],$$
$$[17, 10, 6], [10, 9, 6], [17, 9, 13], [14, 17, 10],$$
$$[1, 9, 10], [6, 10, 17]$$

Here is a sample collection of ten random agents, one of a set of twenty randomly generated ten-agent collections:

$$[19, 18, 7], [11, 14, 8], [13, 10, 15], [12, 13, 5],$$
$$[10, 9, 20], [15, 13, 17], [20, 6, 14], [17, 2, 20],$$
$$[17, 16, 5], [1, 15, 3]$$

Hong and Page introduce "diversity" at this juncture, through an arbitrary definition. Define two ordered triples of integers to have a *diversity rating* lying between 0 and 3, depending on how many of the entries disagree. So $(1, 2, 5)$ and $(1, 4, 5)$ are given a rating of 1, because they disagree in one place, and $(1, 2, 5)$ and $(1, 4, 3)$ get a rating of 2, because they disagree in two places. The second pair of triples is considered "more diverse" than the first. The pair $(1, 2, 3)$ and $(3, 1, 2)$ is even more "diverse," with a diversity rating of the maximal possible 3, since none of the ordered entries match. Adding up the "diversity" of our set of ten random agents over all forty-five possible pairs, we get a total of 131, which is larger than 120, the total "diversity" of our set of ten best agents. With the gentlest of pushes, the *random* group has been recast as the *more diverse* group and the authors make the leap of logic that this group performs better *because* it is more diverse.

This argument has several problems. A misuse of terminology compounds them. First, the authors are apparently unaware of a principle that is widely known in both the theory of probability and the theory of algorithms. This is the idea that randomization can improve algorithms, and often can improve them dramatically. This phenomenon has been studied by mathematicians and computer scientists for forty years. There are many well-known, important algorithms based on this principle, including, for example, primality testing. It is certainly a powerful idea, but not new, and not "diversity."

Second, the authors make the common mistake of confusing correlation with causation. Because the random group had a characteristic to which the authors assigned the name "diversity," they attributed the relative success of the random group to "diversity." However there is no indication that the cause of this random group's success is its "diversity."

Indeed, if its greater "diversity" is really the cause of the group's improved performance, then a group maximizing "diversity" would perform even better than a random group. But our replication of the authors' model shows this is not the case. We ran the simulation with different groups of ten agents that achieved maximal possible "diversity." In all cases, a maximally "diverse" group performed less well than the median performance of 200 random groups of ten agents. In the spirit of [1], we might claim that *randomness trumps diversity*.

This is not unexpected, and it confirms our first point. Not only does randomness help in algorithms, but randomness often does better than any known deterministic procedure. As stated in *Probability and Computing* by Mitzenmacher and Upfal [2], "In…many important applications, randomized algorithms are significantly more efficient than the best known deterministic solutions." The contrived optimization problem in [1] gives an example of a situation where randomly chosen agents perform better than algorithms that choose agents according to deterministic characteristics, whether they are labeled "diversity" or "ability."

Finally, the attempt to assign a standard English meaning to a mathematical phenomenon is fraught with peril. For example, in the "Computational Experiment," instead of giving two ordered triples of integers a diversity rating between zero and three, we could instead assign them a *hostility rating* of between zero and three. Indeed we can do this using precisely the same mathematical definition as before; all we will change is the English word attached to the mathematical definition. One could argue that this is a more natural terminology, since it reflects the extent of disagreement between two triples. Thus we can give $(1, 2, 5)$ and $(1, 4, 5)$ a hostility rating of 1, because they disagree in one place, and $(1, 2, 5)$ and $(1, 4, 3)$ a hostility rating of 2, because they disagree in two places.

What does the simulation show now? It is still the case that the ten random agents acting together outperform the ten best agents acting together. But strikingly, we can observe, using precisely the same information as before, that the random group is much more hostile than the best group. Using Hong and Page's line of reasoning, we would be driven to the conclusion that *hostility trumps ability*. That is, if you are trying to form a team to maximize performance on a task, you should make your selection to maximize mutual antipathy among members of the team. We don't recommend this approach, but it is as well founded as Hong and Page's diversity recommendations.

Summary of Problems

Any one of the problems listed below would be sufficient to invalidate the claims of the authors.

(1) Theorem 1 is incorrect as stated.
(2) Once corrected, Theorem 1 is trivial. It is stated in a way which obscures its meaning. It has no mathematical interest and little content.
(3) Theorem 1 is unrelated to the "Computational Experiment." Not only are the numbers of agents selected too small for the theorem to come into play, the hypotheses of the theorem are generally not met. See the Appendix for a detailed example.
(4) The "Computational Experiment" is a contrived optimization problem in which the restrictions on the algorithms are artificial.
(5) The "Computational Experiment" is an illustration of the benefits of randomness, not "diversity."
(6) The "Computational Experiment" has a simple optimal algorithm; the best algorithm simply checks the value of the function V at every point. That is the "highest ability" algorithm for the problem, and it clearly works better than any other possible combination of alternative algorithms, unless they collectively also always return x^*.
(7) The attempt to equate mathematical quantities with human attributes is inappropriate. For example, to associate two triples of integers $(1, 2, 3)$ and $(3, 1, 2)$ with two "problem-solvers" who have a "diverse" approach to problem-solving is not plausible. It is just as reasonable to say they represent two hostile "problem-solvers."

Who Uses this Result?

The "Diversity Trumps Ability" concept is appealing in certain circles. A Google search for "Diversity Trumps Ability" turns up over 70,000 hits, some referencing [1], and some the book Page wrote on similar themes [3]. Page's work on diversity has been cited by NASA [4], the US Geological Survey [5], and Lawrence Berkeley Labs [6], among many others.

Hong and Page's paper has been used to give a scientific veneer to the diversity field, as it is one of the few research papers that appears to rely on more than qualitative information for its conclusions. Page comments on Theorem 1 in his book *The Difference* [3] (p. 162),

> This theorem is no mere metaphor or cute empirical anecdote that may or may not be true ten years from now. It is a mathematical truth.

This is just wrong. The claim that diversity trumps ability has been given no foundation by Hong and Page's paper.

To summarize, the paper "Groups of diverse problem solvers can outperform groups of high-ability problem solvers" [1] contains a theorem that has neither mathematical content nor real-world applications, and a contrived computer simulation that illustrates the well-known fact that random algorithms are often effective. What the paper emphatically does not contain is information that can be applied to any real-world situation involving actual people.

Acknowledgments

Thanks to Elizabeth Hass for assistance with the computational component of this paper, and for many helpful conversations. I am also grateful to my colleagues who offered helpful advice and comments on drafts of the paper.

Appendix

We provide a counterexample to Theorem 1, and an example to illustrate that the hypotheses of Theorem 1 are not met by the setup of the "Computational Experiment."

For completeness, we restate Theorem 1 as it appears in [1]. We also indicate how to correct the problem demonstrated by the counterexample by adding a hypothesis. Note that Assumption 0 is included in [1] as part of the definition of a "problem-solver."

- Assumption 0: (i) $\forall x \in X, V(\phi(x)) \geq V(x)$. (ii) $\phi(\phi(x)) = \phi(x)$.
- Assumption 1: $\forall \phi \in \Phi, \exists x \in X$ such that $V(\phi(x)) < V(x^*)$; i.e., for each agent, there is some starting point for which V of its stopping point is not the global maximum.
- Assumption 2: $\forall x \in X$, $x \neq x^*$, $\exists \phi \in \Phi$ such that $\phi(x) \neq x$; i.e., no point of X is fixed under all elements of Φ except x^*.
- Assumption 3: Φ has a unique best element.

Theorem 1 of [1]

Let Φ be a group of problem solvers that satisfies assumptions (1)-(3). Let μ be a probability distribution over Φ with full support. Then, with probability one, a sample path will have the following property: there exist positive integers N and N_1, $N \geq N_1$, such that the joint performance of the N_1 independently drawn problem solvers exceeds the joint performance of the N_1 individually best problem solvers among the group of N agents independently drawn from Φ according to μ.

Counter-example to Theorem 1

Let $X = \{a, b, c, d\}$. Define $V(x)$ and three agents ϕ_1, ϕ_2 and ϕ_3 according to the table below:

	a	b	c	d
$V(x)$	1	2	2	3
$\phi_1(x)$	d	b	c	d
$\phi_2(x)$	c	c	c	d
$\phi_3(x)$	b	b	b	d

The set of agents $\Phi = \{\phi_1, \phi_2, \phi_3\}$ satisfies all the hypotheses of Theorem 1. The agents ϕ_1, ϕ_2, ϕ_3 have average values $5/2, 9/4, 9/4$ respectively, so ϕ_1 is the "best" agent. Notice that all three agents acting together do not always return the point d, where the maximum of V occurs. Indeed all three agents acting together work only as well as ϕ_1 acting alone. Hence in this case, no group of agents can outperform ϕ_1, or, equivalently, multiple copies of ϕ_1, hence no N and N_1 exist which satisfy the theorem.

The error occurs in Lemma 1 of [1]. It arises because the informal interpretation of Assumption 2 "When one agent gets stuck, there is always another agent that can find an improvement" ([1], p. 16387) is used; however, this informal interpretation is incorrect. For example, ϕ_1 "gets stuck" at b, however neither ϕ_2 nor ϕ_3 "improves" on $V(b)$. However, as required by Assumption 2, $\phi_2(b) \neq b$.

To avoid this problem, one can add the assumption that V is a one-to-one function.

Theorem 1 and the Computational Experiment

The hypotheses of Theorem 1 are not met by the setup of the "Computational Experiment." Even stipulating that V is a one-to-one function does not correct the problems found here, but we will assume it for convenience. We illustrate that it is highly likely under the setup described that there will be points in X where all the agents "get stuck," violating Assumption 2 of Theorem 1. To see why, consider the case where X consists of the first 2000 integers, and $k = 3$ and $l = 20$, as in the "Computational Experiment." Let x_i be the point where V achieves its maximum. Suppose that x_j is the point where V achieves its next-highest value. Notice that, if $x_i - x_j$ is not between 1 and 20 mod(2000), all agents will get stuck at x_j, violating Assumption 2. Hence it is fairly likely that all agents will get stuck at x_j. This argument can be iterated through the decreasing values of V, making it highly likely that all agents will get stuck at some x for fixed V.

References

[1] L. Hong and S. Page, Groups of diverse problem solvers can outperform groups of high-ability problem solvers, *Proc. Nat. Acad. of Sciences*, November 2004.

[2] Mitzenmacher and Upfal, *Probability and Computing*, Cambridge University Press, 2005.

[3] S. Page, *The Difference*, Princeton University Press, 2007.

[4] http://www.nasa.gov/offices/oce/appel/ask-academy/issues/volume3/AA_3-9_SF_director.html

[5] http://www.usgs.gov/humancapital/documents/actionlearnscenariosresults/2009.0615.ALS-2.Results.pdf

[6] http://today.lbl.gov/2011/05/24